

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number
WO 01/78894 A2

(51) International Patent Classification⁷: **B01L**
(21) International Application Number: PCT/US01/12245
(22) International Filing Date: 13 April 2001 (13.04.2001)
(25) Filing Language: English
(26) Publication Language: English
(30) Priority Data:
09/548,797 13 April 2000 (13.04.2000) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KB, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant: **GENOME THERAPEUTICS CORPORATION** [US/US]; 100 Beaver Street, Waltham, MA 02154 (US).

Published:

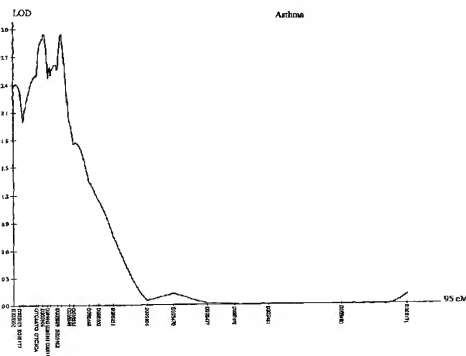
- without international search report and to be republished upon receipt of that report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

(72) Inventor: **KEITH, Tim**; 324 North Road, Bedford, MA 01730 (US).

(74) Agents: **KOMSON, Richard, C.** et al.; Morgan & Finnegan, L.L.P., 345 Park Avenue, New York, NY 10154 (US).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: **NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES, OBESITY, AND INFLAMMATORY BOWEL DISEASE**



(57) Abstract: This invention relates to genes identified from human chromosome 20p13-p12, which are associated with various diseases, including asthma. The invention also relates to the nucleotide sequences of these genes, isolated nucleic acids comprising these nucleotide sequences, and isolated polypeptides or peptides encoded thereby. The invention further relates to vectors and host cells comprising the disclosed nucleotide sequences, or fragments thereof, as well as antibodies that bind to the encoded polypeptides or peptides. Also related are ligands that modulate the activity of the disclosed genes or gene products. In addition, the invention relates to methods and compositions employing the disclosed nucleic acids, polypeptides or peptides, antibodies, and/or ligands for use in diagnostics and therapeutics for asthma and other diseases.

**NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,
OBESITY, AND INFLAMMATORY BOWEL DISEASE**

5

FIELD OF THE INVENTION

10 This invention relates to genes identified from human chromosome
20p13-p12, including Gene 216, which are associated with asthma, obesity,
inflammatory bowel disease, and other human diseases. The invention also
relates to the nucleotide sequences of these genes, including genomic DNA
sequences, cDNA sequences, and single nucleotide polymorphisms. The
15 invention further relates to isolated nucleic acids comprising these nucleotide
sequences, and isolated polypeptides or peptides encoded thereby. Also
related are expression vectors and host cells comprising the disclosed nucleic
acids or fragments thereof, as well as antibodies that bind to the encoded
polypeptides or peptides. The present invention further relates to ligands that
20 modulate the activity of the disclosed genes or gene products. In addition, the
invention relates to diagnostics and therapeutics for various diseases, including
asthma, utilizing the disclosed nucleic acids, polypeptides or peptides,
antibodies, and/or ligands.

BACKGROUND

25 Mouse chromosome 2 has been linked to a variety of disorders including
airway hyperresponsiveness and obesity (DeSanctis et al., 1995, *Nature
Genetics*, 11:150-154; Nagle et al., 1999, *Nature*, 398:148-152). This region
of the mouse genome is homologous to portions of human chromosome 20
including 20p13-p12. Although human chromosome 20p13-12p has been
30 linked to a variety of genetic disorders including diabetes insipidus,
neurohypophyseal, congenital endothelial dystrophy of cornea, insomnia,

neurodegeneration with brain iron accumulation 1 (Hallervorden-Spatz syndrome), fibrodysplasia ossificans progressiva, alagille syndrome, hydrometrocolpos (McKusick-Kaufman syndrome), Creutzfeldt-Jakob disease and Gerstmann-Straussler disease (see NCBI; National Center for
5 Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) the genes affecting these disorders have yet to be discovered. There is a need in the art for identifying specific genes relating to these disorders, as well as genes associated with obesity, lung disease, particularly, inflammatory lung disease
10 phenotypes such as Chronic Obstructive Lung Disease (COPD), Adult Respiratory Distress Syndrome (ARDS), and asthma. Identification and characterization of such genes will make possible the development of effective diagnostics and therapeutic means to treat lung-related disorders.

SUMMARY OF THE INVENTION

15 This invention relates to Gene 216 located on human chromosome 20p13-p12. In specific embodiments, the invention relates to isolated nucleic acids comprising Gene 216 genomic sequences (e.g., SEQ ID NO:5 and SEQ ID NO:6), cDNA sequences (e.g., SEQ ID NO:1 and SEQ ID NO:3), complementary sequences, sequence variants, or fragments thereof, as
20 described herein. The present invention also encompasses nucleic acid probes or primers useful for assaying a biological sample for the presence or expression of Gene 216. The invention further encompasses nucleic acids variants comprising single nucleotide polymorphisms (SNPs) identified in several genes, including Gene 216 (e.g., SEQ ID NO:241-288). Such SNPs
25 can be used to diagnose diseases such as asthma, or to determine a genetic predisposition thereto. In addition, the present invention encompasses nucleic acids comprising alternate splicing variants (e.g., SEQ ID NO:2 and SEQ ID NO:350-362).

This invention also relates to vectors and host cells comprising vectors
30 comprising the Gene 216 nucleic acid sequences disclosed herein. Such vectors can be used for nucleic acid preparations, including antisense nucleic

acids, and for the expression of encoded polypeptides or peptides. Host cells can be prokaryotic or eukaryotic cells. In specific embodiments, an expression vector comprises a DNA sequence encoding the Gene 216 polypeptide sequence (e.g., SEQ ID NO:4 or SEQ ID NO:363), sequence variants, or fragments thereof, as described herein.

The present invention further relates to isolated Gene 216 polypeptides and peptides. In specific embodiments, the polypeptides or peptides comprise the amino acid sequence of the Gene 216 (e.g., SEQ ID NO:4 or SEQ ID NO:363), sequence variants, or portions thereof, as described herein. In addition, this invention encompasses isolated fusion proteins comprising Gene 216 polypeptides or peptides.

The present invention also relates to isolated antibodies, including monoclonal and polyclonal antibodies, and antibody fragments, that are specifically reactive with the Gene 216 polypeptides, fusion proteins, or variants, or portions thereof, as disclosed herein. In specific embodiments, monoclonal antibodies are prepared to be specifically reactive with the Gene 216 polypeptide (e.g., SEQ ID NO:4 or SEQ ID NO:363) or peptides, or sequence variants thereof.

In addition, the present invention relates to methods of obtaining Gene 216 polynucleotides and polypeptides, variant sequences, or fragments thereof, as disclosed herein. Also related are methods of obtaining anti-Gene 216 antibodies and antibody fragments. The present invention also encompasses methods of obtaining Gene 216 ligands, e.g., agonists, antagonists, inhibitors, and binding factors. Such ligands can be used as therapeutics for asthma and related diseases.

The present invention also relates to diagnostic methods and kits utilizing Gene 216 (wild-type, mutant, or variant) nucleic acids, polypeptides, antibodies, or functional fragments thereof. Such factors can be used, for example, in diagnostic methods and kits for measuring expression levels of Gene 216, and to screen for various Gene 216-related diseases, especially asthma. In addition, the nucleic acids described herein can be used to identify

chromosomal abnormalities affecting Gene 216, and to identify allelic variants or mutations of Gene 216 in an individual or population.

The present invention further relates to methods and therapeutics for the treatment of various diseases, including asthma. In various embodiments, therapeutics comprising the disclosed Gene 216 nucleic acids, polypeptides, antibodies, ligands, or variants, derivatives, or portions thereof, are administered to a subject to treat, prevent, or ameliorate asthma. Specifically related are therapeutics comprising Gene 216 antisense nucleic acids, monoclonal antibodies, metalloprotease inhibitors, and gene therapy vectors. Such therapeutics can be administered alone, or in combination with one or more asthma treatments.

In addition, this invention relates to non-human transgenic animals and cell lines comprising one or more of the disclosed Gene 216 nucleic acids, which can be used for drug screening, protein production, and other purposes. Also related are non-human knock-out animals and cell lines, wherein one or more endogenous Gene 216 genes (i.e., orthologs), or portions thereof, are deleted or replaced by marker genes.

This invention further relates to methods of identifying proteins that are candidates for being involved in asthma (i.e., a "candidate protein"). Such proteins are identified by a method comprising: 1) identifying a protein in a first individual having the asthma phenotype; 2) identifying a protein in a second individual not having the asthma phenotype; and 3) comparing the protein of the first individual to the protein of the second individual, wherein a) the protein that is present in the second individual but not the first individual is the candidate protein; or b) the protein that is present in a higher amount in the second individual than in the first individual is the candidate protein; or c) the protein that is present in a lower amount in the second individual than in the first individual is the candidate protein.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 depicts the LOD Plot of Linkage to Asthma.

Figure 2 depicts the LOD Plot of Linkage to BHR (PC20 \leq 4 mg/ml) &

Asthma.

Figure 3 depicts the LOD Plot of Linkage to BHR (PC20 \leq 16 mg/ml) & Asthma

Figure 4 depicts the LOD Plot of Linkage to High Total IgE & Asthma

5 **Figure 5** depicts the LOD Plot of Linkage to High Specific IgE & Asthma

Figure 6 depicts the BAC/STS content contig map of human chromosome 20p13-p12.

Figure 7 depicts the BAC1098L22 nucleotide sequence (SEQ ID NO:5).

10 **Figure 8** depicts the locations of single nucleotide polymorphisms, corresponding amino acid changes, and domains in the Gene 216 transcript. The exons of the transcript are marked from A to T and the size of each one is indicated. Above the exons, the 8 domains are labeled and a black bar represents the approximate location of each one. Underneath the black bars are the approximate location of the amino acid changes that have been
15 identified. The amino acids boxed in white are the alleles that are most frequently observed. The nucleotides boxed in gray are the alleles that are most frequently observed. Single nucleotide polymorphisms are unboxed, and the polymorphism names appear underneath. The uterus cDNA clone does not contain all of Exon A, and does not contain the sequence CAG between
20 Exon S and T.

Figure 9 depicts alternate splice variants of Gene 216 obtained from lung tissue, including rt672 (SEQ ID NO:350), rt690 (SEQ ID NO:351), rt709 (SEQ ID NO:352), rt711 (SEQ ID NO:353), rt713 (SEQ ID NO:354), and rt720 (SEQ ID NO:355).

25 **Figure 10** depicts alternate splice variants of Gene 216 obtained from lung tissue, including rt725 (SEQ ID NO:356), rt727 (SEQ ID NO:357), rt733 (SEQ ID NO:358), rt735 (SEQ ID NO:359), rt764 (SEQ ID NO:360), rt772 (SEQ ID NO:361), and rt774 (SEQ ID NO:362).

Figure 11 depicts the structure of the genomic sequence of Gene 216.

30 **Figure 12** depicts the alternate AG splice sequences at the junction of Intron ST and Exon T in Gene 216.

Figure 13 depicts the promoter region of Gene 216. The Gene 216 promoter sequence is shown in SEQ ID NO:8; the Gene 216 enhancer sequence is shown in SEQ ID NO:7.

Figure 14 depicts a dendrogram of the ADAM family members and the relationship of Gene 216 to ADAMs that possesses an active metalloprotease domain.

Figures 15A-15C depict Northern Blots illustrating Gene 216 expression patterns. Figures 15A-15B show Gene 216 expression in various tissue types. Figure 15C shows Gene 216 expression in bronchial smooth muscle tissue.

Figure 16 depicts a Dot Blot that shows Gene 216 expression in various tissue types.

Figure 17 depicts RT-PCR analysis of Gene 216 expression in primary cells from lung tissue.

Figure 18 depicts an amino acid sequence alignment (Pileup) of 5 ADAM family members that are closely related to Gene 216. Amino acids highlighted in black show 100% identity within the Pileup; dark gray show 80% identity; and light gray show 60% identity. The boxed amino acids represent the cysteine switch, the metalloprotease domain, and the "met-turn". The labeled arrows show the locations of the 8 domains.

Figure 19 depicts the amino acid sequence of Gene 216 (SEQ ID NO:4). Labeled arrows above the sequence denote domain and corresponding length. Black boxes represent the signal sequence and the transmembrane domain identified by hydrophobicity plots. The underlined cysteine residue at position 133 is predicted to be involved in the cysteine switch, the dashed box represents the metalloprotease domain, and the methionine underlined twice is the "met-turn". The gray boxes represent the signaling binding sites identified in the cytoplasmic tail. The amino acid changes corresponding to single nucleotide polymorphisms are indicated in bold. The alanine deleted in the uterus cDNA clone is marked within a black triangle, and if present would have been between the glutamine and the aspartic acid.

Figure 20 depicts the Kyte-Doolittle hydrophobicity plot for the Gene

216 amino acid sequence.

Figures 21 depicts the genomic sequence of the mouse ortholog of Gene 216.(SEQ ID NO:364).

Figure 22 depicts the cDNA nucleotide sequence (SEQ ID NO:364) and
5 predicted amino acid sequence (SEQ ID NO:365) of the mouse ortholog of Gene 216.

Figure 23 depicts an amino acid sequence alignment (Pileup) of human Gene 216 polypeptide (SEQ ID NO:4) and the mouse ortholog of Gene 216 (SEQ ID NO:366). Vertical lines indicate identical amino acid residues. Dots
10 indicate similar amino acid residues.

Figure 24 depicts the nucleotide sequence (SEQ ID NO:1) and encoded amino acid sequence (SEQ ID NO:4) determined from the master cDNA sequence of Gene 216. The master cDNA sequence combines the sequence information from the uterine cDNA clone and 5'RACE clone. Identified single
15 nucleotide polymorphism positions are underlined.

Figure 25 depicts the results of a case control study p-value plot that shows single nucleotide polymorphism association with the asthma phenotype in the combined US and UK populations.

Figure 26 depicts the results of a case control study p-value plot that
20 shows single nucleotide polymorphism association with the asthma phenotype in the US and UK populations, separately.

Figure 27 depicts the results of a case control study p-value plot that shows single nucleotide polymorphism association with the bronchial hyper-responsiveness and asthma phenotypes in the US and UK combined
25 population.

Figure 28 depicts the results of a case control study p-value plot that shows single nucleotide polymorphism association with the bronchial hyper-responsiveness and asthma phenotypes in the US and UK populations, separately.

Figure 29 depicts the genomic nucleotide sequence (SEQ ID NO:6) determined for Gene 216. Identified single nucleotide polymorphism positions
30

are underlined.

Figure 30 depicts the nucleotide sequence (SEQ ID NO:3) and encoded amino acid sequence (SEQ ID NO: 363) of Gene 216 determined from the uterus cDNA clone. Identified single nucleotide polymorphism positions are underlined.

Figure 31 depicts the nucleotide sequence (SEQ ID NO:350) and encoded amino acid sequence (SEQ ID NO:337) of Gene 216 alternate splice variant rt672.

Figure 32 depicts the nucleotide sequence (SEQ ID NO:351) and encoded amino acid sequence (SEQ ID NO:338) of Gene 216 alternate splice variant rt690.

Figure 33 depicts the nucleotide sequence (SEQ ID NO:352) and encoded amino acid sequence (SEQ ID NO:339) of Gene 216 alternate splice variant rt709.

Figure 34 depicts the nucleotide sequence (SEQ ID NO:353) and encoded amino acid sequence (SEQ ID NO:340) of Gene 216 alternate splice variant rt711.

Figure 35 depicts the nucleotide sequence (SEQ ID NO:354) and encoded amino acid sequence (SEQ ID NO:341) of Gene 216 alternate splice variant rt713.

Figure 36 depicts the nucleotide sequence (SEQ ID NO:355) and encoded amino acid sequence (SEQ ID NO:342) of Gene 216 alternate splice variant rt720.

Figure 37 depicts the nucleotide sequence (SEQ ID NO:356) and encoded amino acid sequence (SEQ ID NO:343) of Gene 216 alternate splice variant rt725.

Figure 38 depicts the nucleotide sequence (SEQ ID NO:357) and encoded amino acid sequence (SEQ ID NO:344) of Gene 216 alternate splice variant rt727.

Figure 39 depicts the nucleotide sequence (SEQ ID NO:358) and encoded amino acid sequence (SEQ ID NO:345) of Gene 216 alternate splice

variant rt733.

Figure 40 depicts the nucleotide sequence (SEQ ID NO:359) and encoded amino acid sequence (SEQ ID NO:346) of Gene 216 alternate splice variant rt735.

5 **Figure 41** depicts the nucleotide sequence (SEQ ID NO:360) and encoded amino acid sequence (SEQ ID NO:347) of Gene 216 alternate splice variant rt764.

10 **Figure 42** depicts the nucleotide sequence (SEQ ID NO:361) and encoded amino acid sequence (SEQ ID NO:348) of Gene 216 alternate splice variant rt772.

Figure 43 depicts the nucleotide sequence (SEQ ID NO:362) and encoded amino acid sequence (SEQ ID NO:349) of Gene 216 alternate splice variant rt774.

DETAILED DESCRIPTION OF THE INVENTION

15 Gene 216 was identified by extensive analysis of the region of human chromosome 20p13-p12 associated with airway hyperresponsiveness, asthma, and atopy. This region has also been implicated in other diseases such as obesity (Wilson, 1999, *Arch. Intern. Med.* **159**:2513-4). Bronchial asthma, furthermore, has been linked to intestinal conditions such as inflammatory
20 bowel disease (B. Wallaert et al., 1995, *J. Exp. Med.* **182**:1897-1904). Thus, there was a need to identify and isolate the gene(s) associated with this region of human chromosome 20.

Definitions

25 To aid in the understanding of the specification and claims, the following definitions are provided.

 "Disorder region" refers to a portion of the human chromosome 20 bounded by the markers D20S502 and D20S851. A "disorder-associated" nucleic acid or polypeptide sequence refers to a nucleic acid sequence that maps to region 20p13-p12 or the polypeptides encoded therein (e.g., Gene
30 216 nucleic acids, and polypeptides). For nucleic acids, this encompasses sequences that are identical or complementary to the Gene 216 sequence, as

well as sequence-conservative, function-conservative, and non-conservative variants thereof. For polypeptides, this encompasses sequences that are identical to the Gene 216 polypeptide, as well as function-conservative and non-conservative variants thereof. Included are naturally-occurring mutations of Gene 216 causative of respiratory diseases or obesity, such as but not limited to mutations which cause altered protein levels or stability (e.g., decreased levels, increased levels, expression in an inappropriate tissue type, increased stability, and decreased stability).

As used herein, the "reference sequence" for Gene 216 is BAC1098L22 (SEQ ID NO:5). The BAC1098L22 sequence is also the source of the disclosed Gene 216 genomic sequence (SEQ ID NO:6). "Variant" sequences refer to nucleotide sequences (and the encoded amino acid sequences) that differ from the reference sequence at one or more positions. Non-limiting examples of variant sequences include the disclosed Gene 216 single nucleotide polymorphisms (SNPs), alternate splice variants, and the amino acid sequences encoded by these variants.

"Sequence-conservative" variants are those in which a change of one or more nucleotides in a given codon position results in no alteration in the amino acid encoded at that position (i.e., silent mutations). "Function-conservative" variants are those in which a change in one or more nucleotides in a given codon position results in a polypeptide sequence in which a given amino acid residue in the polypeptide has been replaced by a conservative amino acid substitution as described in detail herein. "Function-conservative" variants also include analogs of a given polypeptide and any polypeptides that have the ability to elicit antibodies specific to a designated polypeptide. "Non-conservative" variants are those in which a change in one or more nucleotides in a given codon position results in a polypeptide sequence in which a given amino acid residue in a polypeptide has been replaced by a non-conservative amino acid substitution as described hereinbelow. "Non-conservative" variants also include polypeptides comprising non-conservative amino acid substitutions.

As used herein, the term "ortholog" denotes a gene or polypeptide obtained from one species that has homology to an analogous gene or polypeptide from a different species. The term "paralog" denotes a gene or polypeptide obtained from a given species that has homology to a distinct gene or polypeptide from that same species. For example, the disclosed mouse and human Gene 216 sequences are orthologs, whereas human Gene 216 and human ADAM 19 are paralogs.

"Nucleic acid or "polynucleotide" as used herein refers to purine- and pyrimidine-containing polymers of any length, either polyribonucleotides or polydeoxyribonucleotide or mixed polyribo-polydeoxyribonucleotides. This includes single- and double-stranded molecules, i.e., DNA-DNA, DNA-RNA and RNA-RNA hybrids, as well as "protein nucleic acids" (PNA) formed by conjugating bases to an amino acid backbone. This also includes nucleic acids containing modified bases.

As used herein, "isolated" nucleic acids are nucleic acids separated away from other components (e.g., DNA, RNA, and protein) with which they are associated (e.g., as obtained from cells, chemical synthesis systems, or phage or nucleic acid libraries). Isolated nucleic acids are at least 60% free, preferably 75% free, and most preferably 90% free from other associated components. In accordance with the present invention, isolated nucleic acids can be obtained by methods described herein, or other established methods, including isolation from natural sources (e.g., cells, tissues, or organs), chemical synthesis, recombinant methods, combinations of recombinant and chemical methods, and library screening methods.

Nucleic acids referred to herein as "recombinant" are nucleic acids which have been produced by recombinant DNA methodology, including those nucleic acids that are generated by procedures which rely upon a method of artificial replication, such as the polymerase chain reaction (PCR) and/or cloning into a vector using restriction enzymes. Portions of recombinant nucleic acids which code for polypeptides can be identified and isolated by, for example, the method of M. Jasin et al., U.S. Patent No. 4,952,501.

A "coding sequence" or a "protein-coding sequence" is a polynucleotide sequence capable of being transcribed into mRNA and/or capable of being translated into a polypeptide or peptide. The boundaries of the coding sequence are typically determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus.

A "complement" of a nucleic acid sequence as used herein refers to the "antisense" sequence that participates in Watson-Crick base-pairing with the original sequence.

A "probe" or "primer" refers to a nucleic acid or oligonucleotide that forms a hybrid structure with a sequence in a target region due to complementarity of the probe or primer sequence to at least one portion of the target region sequence.

Nucleic acids are "hybridizable" to each other when at least one strand of the nucleic acid can anneal to another nucleic acid strand under defined stringency conditions. Hybridization requires that the two nucleic acids contain substantially complementary sequences; depending on the stringency of hybridization, however, mismatches may be tolerated. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementarity, and can be determined in accordance with the methods described herein.

As used herein, "portion" and "fragment" are synonymous. A "portion" as used with regard to a nucleic acid or polynucleotide, refers to fragments of that nucleic acid or polynucleotide. The fragments can range in size from 8 nucleotides to all but one nucleotide of the entire Gene 216 sequence. Preferably, The fragments are at least 8 to 10 nucleotides in length; more preferably at least 12 nucleotides in length; still more preferably at least 15 to 20 nucleotides in length; yet more preferably at least 25 nucleotides in length; and most preferably at least 35 to 55 nucleotides in length.

"cDNA" refers to complementary or copy DNA produced from an RNA template by the action of RNA-dependent DNA polymerase (reverse transcriptase). Thus, a "cDNA clone" means a duplex DNA sequence

complementary to an RNA molecule of interest, included in a cloning vector or PCR amplified. This term includes genes from which the intervening sequences have been removed.

"Cloning" refers to the use of recombination techniques to insert a particular gene or other DNA sequence into a vector molecule. In order to successfully clone a desired gene, it is necessary to use methods for generating DNA fragments, for joining the fragments to vector molecules, for introducing the composite DNA molecule into a host cell in which it can replicate, and for selecting the clone having the target gene from amongst the recipient host cells.

"cDNA library" refers to a collection of recombinant DNA molecules containing cDNA inserts that together comprise essentially all of the expressed genes of an organism. A cDNA library can be prepared by methods known to one skilled in the art (see, e.g., Cowell and Austin, 1997, "cDNA Library Protocols," *Methods in Molecular Biology*). Generally, RNA is first isolated from the cells of the desired organism, and the RNA is used to prepare cDNA molecules.

"Cloning vector" refers to a plasmid or phage DNA or other DNA that is able to replicate in a host cell. The cloning vector is typically characterized by one or more endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without loss of an essential biological function of the DNA, which may contain a marker suitable for use in the identification of cells containing the vector.

"Regulatory sequence" refers to a nucleic acid sequence that controls or regulates expression of structural genes when operably linked to those genes. These include, for example, the lac systems, the trp system, major operator and promoter regions of the phage lambda, the control region of fd coat protein and other sequences known to control the expression of genes in prokaryotic or eukaryotic cells. Regulatory sequences will vary depending on whether the vector is designed to express the operably linked gene in a prokaryotic or eukaryotic host, and may contain transcriptional elements such

as enhancer elements, termination sequences, tissue-specificity elements and/or translational initiation and termination sites.

"Expression vector" refers to a vehicle or plasmid that is capable of expressing a gene that has been cloned into it, after transformation or
5 integration in a host cell. The cloned gene is usually placed under the control of (i.e., operably linked to) a regulatory sequence.

"Operably linked" means that the promoter controls the initiation of expression of the gene. A promoter is operably linked to a sequence of proximal DNA if upon introduction into a host cell the promoter determines the
10 transcription of the proximal DNA sequence(s) into one or more species of RNA. A promoter is operably linked to a DNA sequence if the promoter is capable of initiating transcription of that DNA sequence.

"Host" includes prokaryotes and eukaryotes. The term includes an organism or cell that is the recipient of an expression vector (e.g.,
15 autonomously replicating or integrating vector).

"Amplification" of nucleic acids refers to methods such as polymerase chain reaction (PCR), ligation amplification (or ligase chain reaction, LCR) and amplification methods based on the use of Q-beta replicase. These methods are well known in the art and described, for example, in U.S. Patent Nos.
20 4,683,195 and 4,683,202. Reagents and hardware for conducting PCR are commercially available. Primers useful for amplifying sequences from the disorder region are preferably complementary to, and preferably hybridize specifically to, sequences in the 20p13-p12 region or in regions that flank a target region therein. Gene 216 generated by amplification may be sequenced
25 directly. Alternatively, the amplified sequence(s) may be cloned prior to sequence analysis.

"Gene" refers to a DNA sequence that encodes through its template or messenger RNA a sequence of amino acids characteristic of a specific peptide, polypeptide, or protein. The term "gene" as used herein with reference to
30 genomic DNA includes intervening, non-coding regions, as well as regulatory regions, and can include 5' and 3' ends.

A gene sequence is "wild-type" if such sequence is usually found in individuals unaffected by the disease or condition of interest. However, environmental factors and other genes can also play an important role in the ultimate determination of the disease. In the context of complex diseases
5 involving multiple genes ("oligogenic disease"), the "wild type", or normal sequence can also be associated with a measurable risk or susceptibility, receiving its reference status based on its frequency in the general population. As used herein, "wild-type Gene 216" refers to the reference sequence, BAC1098L22 (SEQ ID NO:5). The wild-type Gene 216 sequence was used to
10 identify the variants (single nucleotide polymorphisms) described in detail herein.

A gene sequence is a "mutant" sequence if it differs from the wild-type sequence. For example, a Gene 216 nucleic acid containing a single nucleotide polymorphism is a mutant sequence. In some cases, the individual
15 carrying such gene has increased susceptibility toward the disease or condition of interest. In other cases, the "mutant" sequence might also refer to a sequence that decreases the susceptibility toward a disease or condition of interest, and thus acting in a protective manner. Also a gene is a "mutant" gene if too much ("overexpressed") or too little ("underexpressed") of such
20 gene is expressed in the tissues in which such gene is normally expressed, thereby causing the disease or condition of interest.

A nucleic acid or fragment thereof is "substantially homologous" to another if, when optimally aligned (with appropriate nucleotide insertions and/or deletions) with the other nucleic acid (or its complementary strand), there is
25 nucleotide sequence identity in at least 60% of the nucleotide bases, usually at least 70%, more usually at least 80%, preferably at least 90%, and more preferably at least 95-98% of the nucleotide bases.

Alternatively, substantial homology exists when a nucleic acid or fragment thereof will hybridize, under selective hybridization conditions, to
30 another nucleic acid (or a complementary strand thereof). Selectivity of hybridization exists when hybridization which is substantially more selective

than total lack of specificity occurs. Typically, selective hybridization will occur when there is at least about 55% sequence identity over a stretch of at least about nine or more nucleotides, preferably at least about 65%, more preferably at least about 75%, and most preferably at least about 90% (M. Kanehisa, 5 1984, *Nucl. Acids Res.* 11:203-213). The length of homology comparison, as described, may be over longer stretches, and in certain embodiments will often be over a stretch of at least 14 nucleotides, usually at least 20 nucleotides, more usually at least 24 nucleotides, typically at least 28 nucleotides, more typically at least 32 nucleotides, and preferably at least 36 or more nucleotides.

10 As used herein, the terms "protein" and "polypeptide" are synonymous. "Peptides" are defined as fragments or portions of polypeptides, preferably fragments or portions having at least one functional activity (e.g., proteolysis, adhesion, fusion, antigenic, or intracellular activity) as the complete polypeptide sequence.

15 "Isolated" polypeptides or peptides are those that are separated from other components (e.g., DNA, RNA, and other polypeptides or peptides) with which they are associated (e.g., as obtained from cells, translation systems, or chemical synthesis systems). In a preferred embodiment, isolated polypeptides or peptides are at least 10% pure; more preferably, 80 or 90% 20 pure. Isolated polypeptides and peptides include those obtained by methods described herein, or other established methods, including isolation from natural sources (e.g., cells, tissues, or organs), chemical synthesis, recombinant methods, or combinations of recombinant and chemical methods. Proteins or polypeptides referred to herein as "recombinant" are proteins or polypeptides 25 produced by the expression of recombinant nucleic acids.

A "portion" as used herein with regard to a protein or polypeptide, refers to fragments of that protein or polypeptide. The fragments can range in size from 5 amino acid residues to all but one residue of the entire protein sequence. Thus, a portion or fragment can be at least 5, 5-50, 50-100, 100- 30 200, 200-400, 400-800, or more consecutive amino acid residues of a Gene 216 protein or polypeptide, for example, SEQ ID NO:4 or SEQ ID NO:363.

An "immunogenic component", is a moiety that is capable of eliciting a humoral and/or cellular immune response in a host animal.

An "antigenic component" is a moiety that binds to its specific antibody with sufficiently high affinity to form a detectable antigen-antibody complex.

5 A "sample" as used herein refers to a biological sample, such as, for example, tissue or fluid isolated from an individual (including, without limitation, plasma, serum, cerebrospinal fluid, lymph, tears, saliva, milk, pus, and tissue exudates and secretions) or from *in vitro* cell culture constituents, as well as samples obtained from, for example, a laboratory procedure.

10 "Antibodies" refer to polyclonal and/or monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof, that can bind to asthma proteins and fragments thereof or to nucleic acid sequences from the 20p13-p12 region, particularly from the asthma locus or a portion thereof.

15 The term antibody is used both to refer to a homogeneous molecular entity, or a mixture such as a serum product made up of a plurality of different molecular entities. Proteins may be prepared synthetically in a protein synthesizer and coupled to a carrier molecule and injected over several months into rabbits. Rabbit sera is tested for immunoreactivity to the protein or fragment. Monoclonal antibodies may be made by injecting mice with the
20 proteins, or fragments thereof. Monoclonal antibodies will be screened by ELISA and tested for specific immunoreactivity with protein or fragments thereof. (Harlow et al., 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). These antibodies will be useful in assays as well as therapeutics.

25 "Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings
30 of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (A.M. Lesk (ed),

- 1988, *Computational Molecular Biology*, Oxford University Press, NY; D.W. Smith (ed), 1993, *Biocomputing. Informatics and Genome Projects*, Academic Press, NY; A.M. Griffin and H.G. Griffin, H. G (eds), 1994, *Computer Analysis of Sequence Data*, Part I, Humana Press, NJ; G. von Heinje, 1987, *Sequence*
5 *Analysis in Molecular Biology*, Academic Press; and M. Gribskov and J. Devereux (eds), 1991, *Sequence Analysis Primer*, M Stockton Press, NY; H. Carillo and D. Lipman, 1988, *SIAM J. Applied Math.*, **48**:1073.

Technical and scientific terms used herein have the meanings commonly understood by one of ordinary skill in the art to which the present
10 invention pertains, unless otherwise defined. Reference is made herein to various methodologies known to those of skill in the art. Publications and other materials setting forth such known methodologies to which reference is made are incorporated herein by reference in their entireties as though set forth in full.

- 15 Standard reference works setting forth the general principles of recombinant DNA technology include J. Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; P.B. Kaufman et al., (eds), 1995, *Handbook of Molecular and Cellular Methods in Biology and Medicine*, CRC Press, Boca
20 Raton; M.J. McPherson (ed), 1991, *Directed Mutagenesis: A Practical Approach*, IRL Press, Oxford; J. Jones, 1992, *Amino Acid and Peptide Synthesis*, Oxford Science Publications, Oxford; B.M. Austen and O.M.R. Westwood, 1991, *Protein Targeting and Secretion*, IRL Press, Oxford; D.N Glover (ed), 1985, *DNA Cloning*, Volumes I and II; M.J. Gait (ed), 1984,
25 *Oligonucleotide Synthesis*; B.D. Hames and S.J. Higgins (eds), 1984, *Nucleic Acid Hybridization*; Wu and Grossman (eds), *Methods in Enzymology* (Academic Press, Inc.), Vol. 154 and Vol. 155; Quirke and Taylor (eds), 1991, *PCR-A Practical Approach*; Hames and Higgins (eds), 1984, *Transcription and Translation*; R.I. Freshney (ed), 1986, *Animal Cell Culture; Immobilized Cells*
30 *and Enzymes*, 1986, IRL Press; Perbal, 1984, *A Practical Guide to Molecular Cloning*; J. H. Miller and M. P. Calos (eds), 1987, *Gene Transfer Vectors for*

- Mammalian Cells*, Cold Spring Harbor Laboratory Press; M.J. Bishop (ed), 1998, *Guide to Human Genome Computing*, 2d Ed., Academic Press, San Diego, CA; L.F. Peruski and A.H. Peruski, 1997, *The Internet and the New Biology: Tools for Genomic and Molecular Research*, American Society for
- 5 Microbiology, Washington, D.C.

- Standard reference works setting forth the general principles of immunology include S. Sell, 1996, *Immunology, Immunopathology & Immunity*, 5th Ed., Appleton & Lange, Publ., Stamford, CT; D. Male et al., 1996, *Advanced Immunology*, 3d Ed., Times Mirror Int'l Publishers Ltd., Publ.,
- 10 London; D.P. Stites and A.I. Terr, 1991, *Basic and Clinical Immunology*, 7th Ed., Appleton & Lange, Publ., Norwalk, CT; and A.K. Abbas et al., 1991, *Cellular and Molecular Immunology*, W. B. Saunders Co., Publ., Philadelphia, PA. Any suitable materials and/or methods known to those of skill can be utilized in carrying out the present invention; however, preferred materials
- 15 and/or methods are described. Materials, reagents, and the like to which reference is made in the following description and examples are generally obtainable from commercial sources, and specific vendors are cited herein.

Nucleic Acids

- The present invention relates to isolated Gene 216 nucleic acids
- 20 comprising genomic DNA within BAC RPCI_1098L22 (e.g., SEQ ID NO:5), the corresponding cDNA sequences (e.g., SEQ ID NO:1 or SEQ ID NO:3), RNA, fragments of the genomic, cDNA, or RNA nucleic acids comprising 20, 40, 60, 100, 200, 500 or more contiguous nucleotides, and the complements thereof.
- Closely related variants are also included as part of this invention, as well as
- 25 nucleic acids sharing at least 50, 60, 70, 80, or 90% identity with the nucleic acids described above, and nucleic acids which would be identical to a Gene 216 nucleic acids except for one or a few substitutions, deletions, or additions.

- The invention also relates to isolated nucleic acids comprising regions required for accurate expression of Gene 216 (e.g., Gene 216 promoter (e.g.,
- 30 SEQ ID NO:8), enhancer (e.g., SEQ ID NO:7), and polyadenylation sequences). In a preferred embodiment, the present invention is directed to

at least 15 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO:1 or SEQ ID NO:6. More particularly, embodiments of this invention include the BAC clone containing segments of Gene 216 including RPCI_1098L22 as set forth in SEQ ID NO:5 (Figure 7).

- 5 The invention further relates to nucleic acids (e.g., DNA or RNA) that hybridize to a) a nucleic acid encoding a Gene 216 polypeptide, such as a nucleic acid having the sequence of SEQ ID NO:1 or SEQ ID NO:6; b) sequence-conservative, function-conservative, and non-conservative variants of (a); and c) fragments or portions of (a) or (b). Nucleic acids that hybridize
- 10 to the sequence of SEQ ID NO:1 or SEQ ID NO:6 can be double- or single-stranded. Hybridization to the sequence of SEQ ID NO:1 or SEQ ID NO:6 includes hybridization to the strand shown or its complementary strand.

- The present invention also relates to nucleic acids that encode a polypeptide having the amino acid sequence of SEQ ID NO:4 or SEQ ID
- 15 NO:363, or functional equivalents thereof. A functional equivalent of a Gene 216 protein includes fragments or variants that perform at least on characteristic function of the Gene 216 protein (e.g., proteolysis, adhesion, fusion, antigenic, or intracellular activity). Preferably, a functional equivalent will share at least 65% sequence identity with the Gene 216 polypeptide.

- 20 In preferred embodiments, nucleic acids of the present invention share at least 50%, preferably at least 60-70%, more preferably at least 70-80% sequence identity, and even more preferably at least 90-100% sequence identity with the sequences of SEQ ID NO:1 or SEQ ID NO:6, or fragments or portions thereof. Sequence identity calculations can be performed using
- 25 computer programs, hybridization methods, or calculations. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package, BLASTN, BLASTX, TBLASTX, and FASTA (J. Devereux et al., 1984, *Nucleic Acids Research* **12**(1):387; S.F. Altschul et al., 1990, *J. Molec. Biol.* **215**:403-
- 30 410; W. Gish and D.J. States, 1994, *Nature Genet.* **3**:266-272; W.R. Pearson and D.J. Lipman, 1988, *Proc Natl. Acad. Sci. USA* **85**(8):2444-8). The BLAST

programs are publicly available from NCBI and other sources . The well-known Smith Waterman algorithm may also be used to determine identity.

For example, nucleotide sequence identity can be determined by comparing a query sequences to sequences in publicly available sequence
5 databases (NCBI) using the BLASTN2 algorithm (S.F. Altschul et al., 1997, *Nucl. Acids Res.*, **25**:3389-3402). The parameters for a typical search are: $E = 0.05$, $v = 50$, $B = 50$, wherein E is the expected probability score cutoff, V is the number of database entries returned in the reporting of the results, and B is the number of sequence alignments returned in the reporting of the results
10 (S.F. Altschul et al., 1990, *J. Mol. Biol.*, **215**:403-410).

In another approach, nucleotide sequence identity can be calculated using the following equation: % identity = (number of identical nucleotides) / (alignment length in nucleotides) * 100. For this calculation, alignment length includes internal gaps but not includes terminal gaps. Alternatively, nucleotide
15 sequence identity can be determined experimentally using the specific hybridization conditions described below.

In accordance with the present invention, polynucleotide alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, insertion, or modification
20 (e.g., via RNA or DNA analogs). Alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. Alterations of a polynucleotide sequence of SEQ ID NO:1 or SEQ
25 ID NO:6 may create nonsense, missense, or frameshift mutations in this coding sequence, and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

Such altered nucleic acids, including DNA or RNA, can be detected and isolated by hybridization under high stringency conditions or moderate
30 stringency conditions, for example, which are chosen to prevent hybridization of nucleic acids having non-complementary sequences. "Stringency

conditions" for hybridizations is a term of art which refers to the conditions of temperature and buffer concentration which permit hybridization of a particular nucleic acid to another nucleic acid in which the first nucleic acid may be perfectly complementary to the second, or the first and second may share
5 some degree of complementarity which is less than perfect.

For example, certain high stringency conditions can be used which distinguish perfectly complementary nucleic acids from those of less complementarity. "High stringency conditions" and "moderate stringency conditions" for nucleic acid hybridizations are explained in F.M. Ausubel et al.
10 (eds), 1995, *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc., New York, NY, the teachings of which are hereby incorporated by reference. In particular, see pages 2.10.1-2.10.16 (especially pages 2.10.8-2.10.11) and pages 6.3.1-6.3.6. The exact conditions which determine the stringency of hybridization depend not only on ionic strength, temperature and the
15 concentration of destabilizing agents such as formamide, but also on factors such as the length of the nucleic acid sequence, base composition, percent mismatch between hybridizing sequences and the frequency of occurrence of subsets of that sequence within other non-identical sequences. Thus, high or moderate stringency conditions can be determined empirically.

20 By varying hybridization conditions from a level of stringency at which no hybridization occurs to a level at which hybridization is first observed, conditions which will allow a given sequence to hybridize with the most similar sequences in the sample can be determined. Preferably the hybridizing sequences will have 60-70% sequence identity, more preferably 70-85%
25 sequence identity, and even more preferably 90-100% sequence identity.

Typically, the hybridization reaction is initially performed under conditions of low stringency, followed by washes of varying, but higher stringency. Reference to hybridization stringency, e.g., high, moderate, or low stringency, typically relates to such washing conditions. Hybridization
30 conditions are based on the melting temperature (T_m) of the nucleic acid probe or primer and are typically classified by degree of stringency of the conditions

under which hybridization is measured (Ausubel et al., 1995). For example, high stringency hybridization typically occurs at about 5-10% C below the T_m ; moderate stringency hybridization occurs at about 10-20% below the T_m ; and low stringency hybridization occurs at about 20-25% below the T_m . The melting temperature can be approximated by the formulas as known in the art, depending on a number of parameters, such as the length of the hybrid or probe in number of nucleotides, or hybridization buffer ingredients and conditions. As a general guide, T_m decreases approximately 1°C with every 1% decrease in sequence identity at any given SSC concentration. Generally, doubling the concentration of SSC results in an increase in T_m of ~17°C. Using these guidelines, the washing temperature can be determined empirically for moderate or low stringency, depending on the level of mismatch sought.

High stringency hybridization conditions are typically carried out at 65 to 68°C in 0.1 X SSC and 0.1% SDS. Highly stringent conditions allow hybridization of nucleic acid molecules having about 95 to 100% sequence identity. Moderate stringency hybridization conditions are typically carried out at 50 to 65°C in 1 X SSC and 0.1% SDS. Moderate stringency conditions allow hybridization of sequences having at least about 80 to 95% nucleotide sequence identity. Low stringency hybridization conditions are typically carried out at 40 to 50°C in 6 X SSC and 0.1% SDS. Low stringency hybridization conditions allow detection of specific hybridization of nucleic acid molecules having at least about 50 to 80% nucleotide sequence identity.

For example, high stringency conditions can be attained by hybridization in 50% formamide, 5 X Denhardt's solution, 5 X SSPE or SSC (1 X SSPE buffer comprises 0.15 M NaCl, 10 mM Na_2HPO_4 , 1 mM EDTA; 1 X SSC buffer comprises 150 mM NaCl, 15 mM sodium citrate, pH 7.0), 0.2% SDS at about 42°C, followed by washing in 1 X SSPE or SSC and 0.1% SDS at a temperature of at least about 42°C, preferably about 55°C, more preferably about 65°C. Moderate stringency conditions can be attained, for example, by hybridization in 50% formamide, 5 X Denhardt's solution, 5 X SSPE or SSC, and 0.2% SDS at 42°C to about 50°C, followed by washing in 0.2 X SSPE or

SSC and 0.2% SDS at a temperature of at least about 42°C, preferably about 55°C, more preferably about 65°C. Low stringency conditions can be attained, for example, by hybridization in 10% formamide, 5 X Denhardt's solution, 6 X SSPE or SSC, and 0.2% SDS at 42°C, followed by washing in 1 X SSPE or
5 SSC, and 0.2% SDS at a temperature of about 45°C, preferably about 50°C in 4 X SSC at 60°C for 30 min.

High stringency hybridization procedures typically (1) employ low ionic strength and high temperature for washing, such as 0.015 M NaCl/ 0.0015 M sodium citrate, pH 7.0 (0.1 X SSC) with 0.1% sodium dodecyl sulfate (SDS) at
10 50°C; (2) employ during hybridization 50% (vol/vol) formamide with 5 X Denhardt's solution (0.1% weight/volume highly purified bovine serum albumin/0.1% wt/vol Ficoll/0.1% wt/vol polyvinylpyrrolidone), 50 mM sodium phosphate buffer at pH 6.5 and 5 X SSC at 42°C; or (3) employ hybridization
15 with 50% formamide, 5 X SSC, 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 X Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 X SSC and 0.1% SDS.

In one particular embodiment, high stringency hybridization conditions may be attained by:

20 -- Prehybridization treatment of the support (e.g. nitrocellulose filter or nylon membrane), to which is bound the nucleic acid capable of hybridizing with any of the sequences of the invention, is carried out at 65°C for 6 hr with a solution having the following composition: 4 X SSC, 10 X Denhardt's (1 X Denhardt's comprises 1% Ficoll, 1% polyvinylpyrrolidone, 1% BSA (bovine serum albumin); 1 X SSC comprises of 0.15 M of NaCl and 0.015 M of sodium citrate, pH 7);
25

-- Replacement of the pre-hybridization solution in contact with the support by a buffer solution having the following composition: 4 X SSC, 1 X Denhardt's, 25 mM NaPO₄, pH 7, 2 mM EDTA, 0.5% SDS, 100 µg/ml of sonicated salmon sperm DNA containing a nucleic acid derived from the
30

sequences of the invention as probe, in particular a radioactive probe, and previously denatured by a treatment at 100°C for 3 min;

-- Incubation for 12 hr at 65°C;

-- Successive washings with the following solutions: 1) four washings with
5 2 X SSC, 1 X Denhardt's, 0.5% SDS for 45 min at 65°C; 2) two washings with
0.2 X SSC, 0.1 X SSC for 45 min at 65°C; and 3) 0.1 x SSC, 0.1% SDS for 45
min at 65°C.

Additional examples of high, medium, and low stringency conditions can
be found in Sambrook et al., 1989. Exemplary conditions are also described
10 in M.H. Krause and S.A. Aaronson, 1991, *Methods in Enzymology*, 200:546-
556; Ausubel et al., 1995. It is to be understood that the low, moderate and
high stringency hybridization/washing conditions may be varied using a variety
of ingredients, buffers, and temperatures well known to and practiced by the
skilled practitioner.

15 Isolated nucleic acids that are characterized by their ability to hybridize
to (a) a nucleic acid encoding a Gene 216 polypeptide, such as the nucleic
acids depicted as SEQ ID NO:1 or SEQ ID NO:6, b) the complement of (a), (c)
or a portion of (a) or (b) (e.g., under high or moderate stringency conditions),
may further encode a protein or polypeptide having at least one function
20 characteristic of a Gene 216 polypeptide, such as proteolysis, adhesion, fusion,
and intracellular activity, or binding of antibodies that also bind to non-
recombinant Gene 216 protein or polypeptide. The catalytic or binding function
of a protein or polypeptide encoded by the hybridizing nucleic acid may be
detected by standard enzymatic assays for activity or binding (e.g., assays that
25 measure the binding of a transit peptide or a precursor, or other components
of the translocation machinery). Enzymatic assays, complementation tests, or
other suitable methods can also be used in procedures for the identification
and/or isolation of nucleic acids which encode a polypeptide having the amino
acid sequence of SEQ ID NO:4 or SEQ ID NO:363, or a functional equivalent
30 of this polypeptide. The antigenic properties of proteins or polypeptides
encoded by hybridizing nucleic acids can be determined by immunological

methods employing antibodies that bind to a Gene 216 polypeptide such as immunoblot, immunoprecipitation and radioimmunoassay. PCR methodology, including RAGE (Rapid Amplification of Genomic DNA Ends), can also be used to screen for and detect the presence of nucleic acids which encode Gene
5 216-like proteins and polypeptides, and to assist in cloning such nucleic acids from genomic DNA. PCR methods for these purposes can be found in M.A. Innis et al., 1990, *PCR Protocols: A Guide to Methods and Applications*, Academic Press, Inc., San Diego, CA., incorporated herein by reference.

It is understood that, as a result of the degeneracy of the genetic code,
10 many nucleic acid sequences are possible which encode a Gene 216-like protein or polypeptide. Some of these will share little identity to the nucleotide sequences of any known or naturally-occurring Gene 216-like gene but can be used to produce the proteins and polypeptides of this invention by selection of combinations of nucleotide triplets based on codon choices. Such variants,
15 while not hybridizable to a naturally-occurring Gene 216 gene under conditions of high stringency, are contemplated within this invention.

Also encompassed by the present invention are alternate splice variants produced by differential processing of the primary transcript(s) from Gene 216 genomic DNA. An alternate splice variant may comprise, for example, the
20 sequence of any one of SEQ ID NO:2 and SEQ ID NO:350-362. Alternate splice variants can also comprise other combinations of introns/exons of SEQ ID NO:1 or SEQ ID NO:6, which can be determined by those of skill in the art.

Alternate splice variants can be determined experimentally, for example, by isolating and analyzing cellular RNAs (e.g., Southern blotting or PCR), or by
25 screening cDNA libraries using the Gene 216 nucleic acid probes or primers described herein. In another approach, alternate splice variants can be predicted using various methods, computer programs, or computer systems available to practitioners in the field.

General methods for splice site prediction can be found in Nakata, 1985,
30 *Nucleic Acids Res.* **13**:5327-5340. In addition, splice sites can be predicted using, for example, the GRAIL™ (E.C. Uberbacher and R.J. Mural, 1991, *Proc.*

- Natl. Acad. Sci. USA*, **88**:11261-11265; E.C. Uberbacher, 1995, *Trends Biotech.*, **13**:497-500; <http://grail.lsd.ornl.gov/grailexp>); GenView (L. Milanesi et al., 1993; *Proceedings of the Second International Conference on Bioinformatics, Supercomputing, and Complex Genome Analysis*, H.A. Lim et al. (eds), World Scientific Publishing, Singapore, pp. 573-588; http://25.itba.mi.cnr.it/~webgene/wwwgene_help.html); SpliceView (<http://www.itba.mi.cnr.it/webgene>); and HSPL (V.V. Solovyev et al., 1994, *Nucleic Acids Res.* **22**:5156-5163; V.V. Solovyev et al., 1994, "The Prediction of Human Exons by Oligonucleotide Composition and Discriminant Analysis of Spliceable Open Reading Frames," R. Altman et al. (eds), *The Second International conference on Intelligent systems for Molecular Biology*, AAAI Press, Menlo Park, CA, pp. 354-362; V.V. Solovyev et al., 1993, "Identification Of Human Gene Functional Regions Based On Oligonucleotide Composition," L. Hunter et al. (eds), *In Proceedings of First International conference on Intelligent System for Molecular Biology*, Bethesda, pp. 371-379) computer systems.

Additionally, computer programs such as GeneParser (E.E. Snyder and G.D. Stormo, 1995, *J. Mol. Biol.* **248**: 1-18; E.E. Snyder and G.D. Stormo, 1993, *Nucl. Acids Res.* **21**(3): 607-613; <http://mcdb.colorado.edu/~eesnyder/GeneParser.html>); MZEF (M.Q. Zhang, 1997, *Proc. Natl. Acad. Sci. USA*, **94**:565-568; <http://argon.cshl.org/genefinder>); MORGAN (S. Salzberg et al., 1998, *J. Comp. Biol.* **5**:667-680; S. Salzberg et al. (eds), 1998, *Computational Methods in Molecular Biology*, Elsevier Science, New York, NY, pp. 187-203); VEIL (J. Henderson et al., 1997, *J. Comp. Biol.* **4**:127-141); GeneScan (S. Tiwari et al., 1997, *CABIOS (Bioinformatics)* **13**: 263-270); GeneBuilder (L. Milanesi et al., 1999, *Bioinformatics* **15**:612-621); Eukaryotic GeneMark (J. Besemer et al., 1999, *Nucl. Acids Res.* **27**:3911-3920); and FEXH (V.V. Solovyev et al., 1994, *Nucleic Acids Res.* **22**:5156-5163). In addition, splice sites (i.e., former or potential splice sites) in cDNA sequences can be predicted using, for example, the RNASPL (V.V. Solovyev et al., 1994, *Nucleic Acids Res.* **22**:5156-5163); or INTRON (A. Globek et al., 1991, INTRON version 1.1 manual, Laboratory of Biochemical Genetics, NIMH, Washington, D.C.)

programs.

The present invention also encompasses naturally-occurring polymorphisms of Gene 216. As will be understood by those in the art, the genomes of all organisms undergo spontaneous mutation in the course of their continuing evolution generating variant forms of gene sequences (Gusella, 5 1986, *Ann. Rev. Biochem.* **55**:831-854). Restriction fragment length polymorphisms (RFLPs) include variations in DNA sequences that alter the length of a restriction fragment in the sequence (Botstein et al., 1980, *Am. J. Hum. Genet.* **32**, 314-331 (1980). RFLPs have been widely used in human and animal genetic analyses (see WO 90/13668; WO90/11369; Donis-Keller, 10 1987, *Cell* **51**:319-337; Lander et al., 1989, *Genetics* **121**: 85-99). Short tandem repeats (STRs) include tandem di-, tri- and tetranucleotide repeated motifs, also termed variable number tandem repeat (VNTR) polymorphisms. VNTRs have been used in identity and paternity analysis (U.S. Pat. No. 15 5,075,217; Armour et al., 1992, *FEBS Lett.* **307**:113-115; Horn et al., WO 91/14003; Jeffreys, EP 370,719), and in a large number of genetic mapping studies.

Single nucleotide polymorphisms (SNPs) are far more frequent than RFLPS, STRs, and VNTRs. SNPs may occur in protein coding (e.g., exon), or 20 non-coding (e.g., intron, 5'UTR, 3'UTR) sequences. SNPs in protein coding regions may comprise silent mutations that do not alter the amino acid sequence of a protein. Alternatively, SNPs in protein coding regions may produce conservative or non-conservative amino acid changes, described in detail below. In some cases, SNPs may give rise to the expression of a defective or other variant protein and, potentially, a genetic disease. SNPs 25 within protein-coding sequences can give rise to genetic diseases, for example, in the β -globin (sickle cell anemia) and CFTR (cystic fibrosis) genes. In non-coding sequences, SNPs may also result in defective protein expression (e.g., as a result of defective splicing). Other single nucleotide polymorphisms have 30 no phenotypic effects.

Single nucleotide polymorphisms can be used in the same manner as

RFLPs and VNTRs, but offer several advantages. Single nucleotide polymorphisms tend to occur with greater frequency and are typically spaced more uniformly throughout the genome than other polymorphisms. Also, different SNPs are often easier to distinguish than other types of polymorphisms (e.g., by use of assays employing allele-specific hybridization probes or primers). In one embodiment of the present invention, a Gene 216 nucleic acid contains at least one SNP as set forth in Table 10, herein below. Various combinations of these SNPs are also encompassed by the invention. In a preferred aspect, a Gene 216 SNP is associated with a lung-related disorder, such as asthma.

The nucleic acid sequences of the present invention may be derived from a variety of sources including DNA, cDNA, synthetic DNA, synthetic RNA, or combinations thereof. Such sequences may comprise genomic DNA, which may or may not include naturally occurring introns. Moreover, such genomic DNA may be obtained in association with promoter regions or poly (A) sequences. The sequences, genomic DNA, or cDNA may be obtained in any of several ways. Genomic DNA can be extracted and purified from suitable cells by means well known in the art. Alternatively, mRNA can be isolated from a cell and used to produce cDNA by reverse transcription or other means.

The nucleic acids described herein are used in the methods of the present invention for production of proteins or polypeptides, through incorporation into cells, tissues, or organisms. In one embodiment, DNA containing all or part of the coding sequence for a Gene 216 polypeptide, or DNA which hybridizes to DNA having the sequence SEQ ID NO:1 or SEQ ID NO:6, is incorporated into a vector for expression of the encoded polypeptide in suitable host cells. The encoded polypeptide consisting of Gene 216, or its functional equivalent is capable of normal activity, such as proteolysis, adhesion, fusion, and intracellular activity.

The invention also concerns the use of the nucleotide sequence of the nucleic acids of this invention to identify DNA probes for Gene 216 genes, PCR primers to amplify Gene 216 genes, nucleotide polymorphisms in Gene 216

genes, and regulatory elements of the Gene 216 genes.

The nucleic acids of the present invention find use as primers and templates for the recombinant production of disorder-associated peptides or polypeptides, for chromosome and gene mapping, to provide antisense
5 sequences, for tissue distribution studies, to locate and obtain full length genes, to identify and obtain homologous sequences (wild-type and mutants), and in diagnostic applications.

Probes may also be used for the detection of Gene 216-related sequences, and should preferably contain at least 50%, preferably at least
10 80%, identity to Gene 216 polynucleotide, or a complementary sequence, or fragments thereof. The probes of this invention may be DNA or RNA, the probes may comprise all or a portion of the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:6, or a complementary sequence thereof, and may include
15 promoter, enhancer elements, and introns of the naturally occurring Gene 216 polynucleotide.

The probes and primers based on the Gene 216 gene sequences disclosed herein are used to identify homologous Gene 216 gene sequences and proteins in other species. These Gene 216 gene sequences and proteins are used in the *diagnostic/prognostic, therapeutic and drug-screening methods*
20 described herein for the species from which they have been isolated.

Vectors and Host Cells

The invention also provides vectors comprising the disorder-associated sequences, or derivatives or fragments thereof, and host cells for the production of purified proteins. A large number of vectors, including bacterial,
25 yeast, and mammalian vectors, have been described for replication and/or expression in various host cells or cell-free systems, and may be used for gene therapy as well as for simple cloning or protein expression.

In one aspect, an expression vectors comprises a nucleic acid encoding a Gene 216 polypeptide or peptide, as described herein, operably linked to at
30 least one regulatory sequence. Regulatory sequences are known in the art and are selected to direct expression of the desired protein in an appropriate

host cell. Accordingly, the term regulatory sequence includes promoters, enhancers and other expression control elements (see D.V. Goeddel (1990) *Methods Enzymol.* **185**:3-7). Enhancer and other expression control sequences are described in *Enhancers and Eukaryotic Gene Expression*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1983). It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transfected and/or the type of polypeptide desired to be expressed.

Several regulatory elements (e.g., promoters) have been isolated and shown to be effective in the transcription and translation of heterologous proteins in the various hosts. Such regulatory regions, methods of isolation, manner of manipulation, etc. are known in the art. Non-limiting examples of bacterial promoters include the β -lactamase (penicillinase) promoter; lactose promoter; tryptophan (trp) promoter; araBAD (arabinose) operon promoter; lambda-derived P₁ promoter and N gene ribosome binding site; and the hybrid tac promoter derived from sequences of the trp and lac UV5 promoters. Non-limiting examples of yeast promoters include the 3-phosphoglycerate kinase promoter, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter, galactokinase (GAL1) promoter, galactose epimerase promoter, and alcohol dehydrogenase (ADH1) promoter. Suitable promoters for mammalian cells include, without limitation, viral promoters, such as those from Simian Virus 40 (SV40), Rous sarcoma virus (RSV), adenovirus (ADV), and bovine papilloma virus (BPV). Preferred replication and inheritance systems include M13, ColE1, SV40, baculovirus, lambda, adenovirus, CEN ARS, 2 μ m ARS and the like. While expression vectors may replicate autonomously, they may also replicate by being inserted into the genome of the host cell, by methods well known in the art.

To obtain expression in eukaryotic cells, terminator sequences, polyadenylation sequences, and enhancer sequences that modulate gene expression may be required. Sequences that cause amplification of the gene may also be desirable. These sequences are well known in the art.

Furthermore, sequences that facilitate secretion of the recombinant product from cells, including, but not limited to, bacteria, yeast, and animal cells, such as secretory signal sequences and/or preprotein or proprotein sequences, may also be included. Such sequences are well described in the art.

- 5 Expression and cloning vectors will likely contain a selectable marker, a gene encoding a protein necessary for survival or growth of a host cell transformed with the vector. The presence of this gene ensures growth of only those host cells that express the inserts. Typical selection genes encode proteins that 1) confer resistance to antibiotics or other toxic substances, e.g.
- 10 ampicillin, neomycin, methotrexate, etc.; 2) complement auxotrophic deficiencies, or 3) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for Bacilli. Markers may be an inducible or non-inducible gene and will generally allow for positive selection.
- 15 Non-limiting examples of markers include the ampicillin resistance marker (i.e., beta-lactamase), tetracycline resistance marker, neomycin/kanamycin resistance marker (i.e., neomycin phosphotransferase), dihydrofolate reductase, glutamine synthetase, and the like. The choice of the proper selectable marker will depend on the host cell, and appropriate markers for different hosts as understood by those of skill in the art.

- 20 Suitable expression vectors for use with the present invention include, but are not limited to, pUC, pBluescript (Stratagene), pET (Novagen, Inc., Madison, WI), and pREP (Invitrogen) plasmids. Vectors can contain one or more replication and inheritance systems for cloning or expression, one or more markers for selection in the host, e.g. antibiotic resistance, and one or
- 25 more expression cassettes. The inserted coding sequences can be synthesized by standard methods, isolated from natural sources, or prepared as hybrids. Ligation of the coding sequences to transcriptional regulatory elements (e.g., promoters, enhancers, and/or insulators) and/or to other amino acid encoding sequences can be carried out using established methods.

- 30 Suitable cell-free expression systems for use with the present invention include, without limitation, rabbit reticulocyte lysate, wheat germ extract, canine

pancreatic microsomal membranes, *E. coli* S30 extract, and coupled transcription/translation systems (Promega Corp., Madison, WI). These systems allow the expression of recombinant polypeptides or peptides upon the addition of cloning vectors, DNA fragments, or RNA sequences containing

5 protein-coding regions and appropriate promoter elements.

Non-limiting examples of suitable host cells include bacteria, archaea, insect, fungi (e.g., yeast), plant, and animal cells (e.g., mammalian, especially human). Of particular interest are *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, SF9 cells, C129 cells, 293 cells, *Neurospora*, and

10 immortalized mammalian myeloid and lymphoid cell lines. Techniques for the propagation of mammalian cells in culture are well-known (see, Jakoby and Pastan (eds), 1979, *Cell Culture. Methods in Enzymology*, volume 58, Academic Press, Inc., Harcourt Brace Jovanovich, NY). Examples of commonly used mammalian host cell lines are VERO and HeLa cells, CHO

15 cells, and WI38, BHK, and COS cell lines, although it will be appreciated by the skilled practitioner that other cell lines may be used, e.g., to provide higher expression desirable glycosylation patterns, or other features.

Host cells can be transformed, transfected, or infected as appropriate by any suitable method including electroporation, calcium chloride-, lithium chloride-, lithium acetate/polyethylene glycol-, calcium phosphate-, DEAE-dextran-, liposome-mediated DNA uptake, spheroplasting, injection, microinjection, microprojectile bombardment, phage infection, viral infection, or other established methods. Alternatively, vectors containing the nucleic acids of interest can be transcribed *in vitro*, and the resulting RNA introduced

20 into the host cell by well-known methods, e.g., by injection (see, Kubo et al., 1988, *FEBS Letts.* **241**:119). The cells into which have been introduced nucleic acids described above are meant to also include the progeny of such cells.

The nucleic acids of the invention may be isolated directly from cells.

30 Alternatively, the polymerase chain reaction (PCR) method can be used to produce the nucleic acids of the invention, using either RNA (e.g., mRNA) or

DNA (e.g., genomic DNA) as templates. Primers used for PCR can be synthesized using the sequence information provided herein and can further be designed to introduce appropriate new restriction sites, if desirable, to facilitate incorporation into a given vector for recombinant expression.

- 5 Using the information provided in SEQ ID NO:1 and SEQ ID NO:6, one skilled in the art will be able to clone and sequence all representative nucleic acids of interest, including nucleic acids encoding complete protein-coding sequences. It is to be understood that non-protein-coding sequences contained within SEQ ID NO:1 and SEQ ID NO:3 and the genomic sequences
10 of SEQ ID NO:6 and SEQ ID NO:5 are also within the scope of the invention. Such sequences include, without limitation, sequences important for replication, recombination, transcription, and translation. Non-limiting examples include promoters and regulatory binding sites involved in regulation of gene expression, and 5'- and 3'- untranslated sequences (e.g., ribosome-
15 binding sites) that form part of mRNA molecules.

- The nucleic acids of this invention can be produced in large quantities by replication in a suitable host cell. Natural or synthetic nucleic acid fragments, comprising at least ten contiguous bases coding for a desired peptide or polypeptide can be incorporated into recombinant nucleic acid
20 constructs, usually DNA constructs, capable of introduction into and replication in a prokaryotic or eukaryotic cell. Usually the nucleic acid constructs will be suitable for replication in a unicellular host, such as yeast or bacteria, but may also be intended for introduction to (with and without integration within the genome) cultured mammalian or plant or other eukaryotic cells, cell lines,
25 tissues, or organisms. The purification of nucleic acids produced by the methods of the present invention is described, for example, in Sambrook et al., 1989; F.M. Ausubel et al., 1992, *Current Protocols in Molecular Biology*, J. Wiley and Sons, New York, NY.

- The nucleic acids of the present invention can also be produced by
30 chemical synthesis, e.g., by the phosphoramidite method described by Beaucage et al., 1981, *Tetra. Letts.* 22:1859-1862, or the triester method

according to Matteucci et al., 1981, *J. Am. Chem. Soc.*, **103**:3185, and can performed on commercial, automated oligonucleotide synthesizers. A double-stranded fragment may be obtained from the single-stranded product of chemical synthesis either by synthesizing the complementary strand and
5 annealing the strands together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

These nucleic acids can encode full-length variant forms of proteins as well as the wild-type protein. The variant proteins (which could be especially
10 useful for detection and treatment of disorders) will have the variant amino acid sequences encoded by the polymorphisms described in Table 10, when said polymorphisms are read so as to be in-frame with the full-length coding sequence of which it is a component.

Large quantities of the nucleic acids and proteins of the present
15 invention may be prepared by expressing the Gene 216 nucleic acids or portions thereof in vectors or other expression vehicles in compatible prokaryotic or eukaryotic host cells. The most commonly used prokaryotic hosts are strains of *Escherichia coli*, although other prokaryotes, such as *Bacillus subtilis* or *Pseudomonas* may also be used. Mammalian or other
20 eukaryotic host cells, such as those of yeast, filamentous fungi, plant, insect, or amphibian or avian species, may also be useful for production of the proteins of the present invention. For example, insect cell systems (i.e., lepidopteran host cells and baculovirus expression vectors) are particularly suited for large-scale protein production.

25 Host cells carrying an expression vector (i.e., transformants or clones) are selected using markers depending on the mode of the vector construction. The marker may be on the same or a different DNA molecule, preferably the same DNA molecule. In prokaryotic hosts, the transformant may be selected, e.g., by resistance to ampicillin, tetracycline or other antibiotics. Production of
30 a particular product based on temperature sensitivity may also serve as an appropriate marker.

Prokaryotic or eukaryotic cells comprising the nucleic acids of the present invention will be useful not only for the production of the nucleic acids and proteins of the present invention, but also, for example, in studying the characteristics of Gene 216 proteins. Cells and animals that carry the Gene 216 gene can be used as model systems to study and test for substances that have potential as therapeutic agents. The cells are typically cultured mesenchymal stem cells. These may be isolated from individuals with somatic or germline Gene 216 gene. Alternatively, the cell line can be engineered to carry the Gene 216 genes, as described above. After a test substance is applied to the cells, the transformed phenotype of the cell is determined. Any trait of transformed cells can be assessed, including respiratory diseases including asthma, atopy, and response to application of putative therapeutic agents.

Antisense Nucleic Acids

A further embodiment of the invention is antisense nucleic acids or oligonucleotides that are complementary, in whole or in part, to a target molecule comprising a sense strand of Gene 216. The Gene 216 target can be DNA, or its RNA counterpart (i.e., wherein thymine (T) is present in DNA and uracil (U) is present in RNA). When introduced into a cell, antisense nucleic acids or oligonucleotides can hybridize to all or a part of the sense strand of Gene 216, thereby inhibiting gene expression or replication.

In a particular embodiment of the invention, an antisense nucleic acid or oligonucleotide is wholly or partially complementary to, and can hybridize with, a target nucleic acid (either DNA or RNA) having the sequence of SEQ ID NO:1 or SEQ ID NO:6. For example, an antisense nucleic acid or oligonucleotide comprising 16 nucleotides can be sufficient to inhibit expression of the Gene 216 protein. Alternatively, an antisense nucleic acid or oligonucleotide can be complementary to 5' or 3' untranslated regions, or can overlap the translation initiation codon (5' untranslated and translated regions) of the Gene 216 gene, or its functional equivalent. In another embodiment, the antisense nucleic acid is wholly or partially complementary

to, and can hybridize with, a target nucleic acid that encodes a Gene 216 polypeptide.

In addition, oligonucleotides can be constructed which will bind to duplex nucleic acid (i.e., DNA:DNA or DNA:RNA), to form a stable triple helix-
5 containing or triplex nucleic acid. Such triplex oligonucleotides can inhibit transcription and/or expression of a gene encoding Gene 216, or its functional equivalent (M.D. Frank-Kamenetskii and S.M. Mirkin, 1995, *Ann. Rev. Biochem.* **64**:65-95). Triplex oligonucleotides are constructed using the base-pairing rules of triple helix formation and the nucleotide sequence of the gene
10 or mRNA for Gene 216.

The present invention encompasses methods of using oligonucleotides in antisense inhibition of the function of Gene 216. In the context of this invention, the term "oligonucleotide" refers to naturally-occurring species or
15 synthetic species formed from naturally-occurring subunits or their close homologs. The term may also refer to moieties that function similarly to oligonucleotides, but have non-naturally-occurring portions. Thus, oligonucleotides may have altered sugar moieties or inter-sugar linkages. Exemplary among these are phosphorothioate and other sulfur containing species which are known in the art.

In preferred embodiments, at least one of the phosphodiester bonds of the oligonucleotide has been substituted with a structure that functions to enhance the ability of the compositions to penetrate into the region of cells where the RNA whose activity is to be modulated is located. It is preferred that such substitutions comprise phosphorothioate bonds, methyl phosphonate
25 bonds, or short chain alkyl or cycloalkyl structures. In accordance with other preferred embodiments, the phosphodiester bonds are substituted with structures which are, at once, substantially non-ionic and non-chiral, or with structures which are chiral and enantiomerically specific. Persons of ordinary skill in the art will be able to select other linkages for use in the practice of the
30 invention.

Oligonucleotides may also include species that include at least some

modified base forms. Thus, purines and pyrimidines other than those normally found in nature may be so employed. Similarly, modifications on the furanosyl portions of the nucleotide subunits may also be effected, as long as the essential tenets of this invention are adhered to. Examples of such modifications are 2'-O-alkyl- and 2'-halogen-substituted nucleotides. Some non-limiting examples of modifications at the 2' position of sugar moieties which are useful in the present invention include OH, SH, SCH₃, F, OCH₃, OCN, O(CH₂)_n NH₂ and O(CH₂)_n CH₃, where n is from 1 to about 10. Such oligonucleotides are functionally interchangeable with natural oligonucleotides or synthesized oligonucleotides, which have one or more differences from the natural structure. All such analogs are comprehended by this invention so long as they function effectively to hybridize with Gene 216 DNA or RNA to inhibit the function thereof.

The oligonucleotides in accordance with this invention preferably comprise from about 3 to about 50 subunits. It is more preferred that such oligonucleotides and analogs comprise from about 8 to about 25 subunits and still more preferred to have from about 12 to about 20 subunits. As defined herein, a "subunit" is a base and sugar combination suitably bound to adjacent subunits through phosphodiester or other bonds.

Antisense nucleic acids or oligonucleotides can be produced by standard techniques (see, e.g., Shewmaker et al., U.S. Patent No. 5,107,065. The oligonucleotides used in accordance with this invention may be conveniently and routinely made through the well-known technique of solid phase synthesis. Equipment for such synthesis is available from several vendors, including PE Applied Biosystems (Foster City, CA). Any other means for such synthesis may also be employed, however, the actual synthesis of the oligonucleotides is well within the abilities of the practitioner. It is also well known to prepare other oligonucleotide such as phosphorothioates and alkylated derivatives.

The oligonucleotides of this invention are designed to be hybridizable with Gene 216 RNA (e.g., mRNA) or DNA. For example, an oligonucleotide

(e.g., DNA oligonucleotide) that hybridizes to Gene 216 mRNA can be used to target the mRNA for RnaseH digestion. Alternatively, an oligonucleotide that hybridizes to the translation initiation site of Gene 216 mRNA can be used to prevent translation of the mRNA. In another approach, oligonucleotides that
5 bind to the double-stranded DNA of Gene 216 can be administered. Such oligonucleotides can form a triplex construct and inhibit the transcription of the DNA encoding Gene 216 polypeptides. Triple helix pairing prevents the double helix from opening sufficiently to allow the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex
10 DNA have been described (see, e.g., J.E. Gee et al., 1994, *Molecular and Immunologic Approaches*, Futura Publishing Co., Mt. Kisco, NY).

As non-limiting examples, antisense oligonucleotides may be targeted to hybridize to the following regions: mRNA cap region; translation initiation site; translational termination site; transcription initiation site; transcription
15 termination site; polyadenylation signal; 3' untranslated region; 5' untranslated region; 5' coding region; mid coding region; and 3' coding region. Preferably, the complementary oligonucleotide is designed to hybridize to the most unique 5' sequence Gene 216, including any of about 15-35 nucleotides spanning the 5' coding sequence. Appropriate oligonucleotides can be designed using
20 OLIGO software (Molecular Biology Insights, Inc., Cascade, CO; <http://www.oligo.net>).

In accordance with the present invention, the antisense oligonucleotide can be synthesized, formulated as a pharmaceutical composition, and administered to a subject. The synthesis and utilization of antisense and triplex
25 oligonucleotides have been previously described (e.g., H. Simon et al., 1999, *Antisense Nucleic Acid Drug Dev.* **9**:527-31; F.X. Barre et al., 2000, *Proc. Natl. Acad. Sci. USA* **97**:3084-3088; R. Elez et al., 2000, *Biochem. Biophys. Res. Commun.* **269**:352-6; E.R. Sauter et al., 2000, *Clin. Cancer Res.* **6**:654-60). Alternatively, expression vectors derived from retroviruses, adenovirus, herpes
30 or vaccinia viruses, or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population.

Methods which are well known to those skilled in the art can be used to construct recombinant vectors which will express nucleic acid sequence that is complementary to the nucleic acid sequence encoding a Gene 216 polypeptide. These techniques are described both in Sambrook et al., 1989
5 and in Ausubel et al., 1992. For example, Gene 216 expression can be inhibited by transforming a cell or tissue with an expression vector that expresses high levels of untranslatable sense or antisense Gene 216 sequences. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous
10 nucleases. Transient expression may last for a month or more with a non-replicating vector, and even longer if appropriate replication elements included in the vector system.

Various assays may be used to test the ability of Gene 216-specific antisense oligonucleotides to inhibit Gene 216 expression. For example, Gene
15 216 mRNA levels can be assessed northern blot analysis (Sambrook et al., 1989; Ausubel et al., 1992; J.C. Alwine et al. 1977, *Proc. Natl. Acad. Sci. USA* **74**:5350-5354; I.M. Bird, 1998, *Methods Mol. Biol.* **105**:325-36), quantitative or semi-quantitative RT-PCR analysis (see, e.g., W.M. Freeman et al., 1999, *Biotechniques* **26**:112-122; Ren et al., 1998, *Mol. Brain Res.* **59**:256-63; J.M.
20 Cale et al., 1998, *Methods Mol. Biol.* **105**:351-71), or *in situ* hybridization (reviewed by A.K. Raap, 1998, *Mutat. Res.* **400**:287-298). Alternatively, antisense oligonucleotides may be assessed by measuring levels of Gene 216 polypeptide, e.g., by western blot analysis, indirect immunofluorescence, immunoprecipitation techniques (see, e.g., J.M. Walker, 1998, *Protein*
25 *Protocols on CD-ROM*, Humana Press, Totowa, NJ).

Polypeptides

The invention also relates to polypeptides and peptides encoded by the novel nucleic acids described herein. The polypeptides and peptides of this invention can be isolated and/or recombinant. In a preferred embodiment, the
30 Gene 216 polypeptide, or analog or portion thereof, has at least one function characteristic of a Gene 216 protein, for example, proteolysis, adhesion,

fusion, antigenic, and intracellular activity. Protein analogs include, for example, naturally-occurring or genetically engineered Gene 216 variants (e.g. mutants) and portions thereof. Variants may differ from wild-type Gene 216 protein by the addition, deletion, or substitution of one or more amino acid residues. In specific embodiments, polypeptide variants are encoded by Gene 216 nucleic acids containing one or more of the SNPs disclosed herein. Variants also include polypeptides in which one or more residues are modified (i.e., by phosphorylation, sulfation, acylation, etc.), and mutants comprising one or more modified residues.

Variant polypeptides can have conservative changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More infrequently, a variant polypeptide can have non-conservative changes, e.g., substitution of a glycine with a tryptophan. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity can be found using computer programs well known in the art, for example, DNASTAR software (DNASTAR, Inc., Madison, WI)

As non-limiting examples, conservative substitutions in the Gene 216 amino acid sequence can be made in accordance with the following table:

20

Original Residue	Conservative Substitution(s)
Ala	Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	Met, Leu, Tyr
Ser	Thr
Thr	Ser

Trp	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

Substantial changes in function or immunogenicity can be made by selecting substitutions that are less conservative than those shown in the table, above. For example, non-conservative substitutions can be made which more significantly affect the structure of the polypeptide in the area of the alteration, for example, the alpha-helical, or beta-sheet structure; the charge or hydrophobicity of the molecule at the target site; or the bulk of the side chain. The substitutions which generally are expected to produce the greatest changes in the polypeptide's properties are those where 1) a hydrophilic residue, e.g., seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl, or alanyl; 2) a cysteine or proline is substituted for (or by) any other residue; 3) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or 4) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) a residue that does not have a side chain, e.g., glycine.

In one embodiment, polypeptides of the present invention share at least 50% amino acid sequence identity with a Gene 216 polypeptide, such as SEQ ID NO:4, or fragments thereof. Preferably, the polypeptides share at least 65% amino acid sequence identity; more preferably, the polypeptides share at least 75% amino acid sequence identity; even more preferably, the polypeptides share at least 80% amino acid sequence identity with a Gene 216 polypeptide; still more preferably the polypeptides share at least 90% amino acid sequence identity with a Gene 216 polypeptide.

Percent sequence identity can be calculated using computer programs or direct sequence comparison. Preferred computer program methods to determine identity between two sequences include, but are not limited to, the GCG program package, FASTA, BLASTP, and TBLASTN (see, e.g., D.W. Mount, 2001, *Bioinformatics: Sequence and Genome Analysis*, Cold Spring

Harbor Laboratory Press, Cold Spring Harbor, NY). The BLASTP and TBLASTN programs are publicly available from NCBI and other sources. The well-known Smith Waterman algorithm may also be used to determine identity.

Exemplary parameters for amino acid sequence comparison include the following: 1) algorithm from Needleman and Wunsch, 1970, *J Mol. Biol.* 48:443-453; 2) BLOSSUM62 comparison matrix from Hentikoff and Hentikoff, 1992, *Proc. Natl. Acad. Sci. USA* 89:10915-10919; 3) gap penalty = 12; and 4) gap length penalty = 4. A program useful with these parameters is publicly available as the "gap" program (Genetics Computer Group, Madison, WI). The aforementioned parameters are the default parameters for polypeptide comparisons (with no penalty for end gaps).

Alternatively, polypeptide sequence identity can be calculated using the following equation: % identity = (the number of identical residues) / (alignment length in amino acid residues) * 100. For this calculation, alignment length includes internal gaps but does not include terminal gaps.

In accordance with the present invention, polypeptide sequences may be identical to the sequence of SEQ ID NO:4, or may include up to a certain integer number of amino acid alterations. Polypeptide alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. Alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. In specific embodiments, polypeptide variants may be encoded by Gene 216 nucleic acids comprising SNPs and/or alternate splice variants.

The invention also relates to isolated, synthesized and/or recombinant portions or fragments of a Gene 216 protein or polypeptide as described herein. Polypeptide fragments (i.e., peptides) can be made which have full or partial function on their own, or which when mixed together (though fully, partially, or nonfunctional alone), spontaneously assemble with one or more

other polypeptides to reconstitute a functional protein having at least one functional characteristic of a Gene 216 protein of this invention. In addition, Gene 216 polypeptide fragments may comprise, for example, one or more domains of the Gene 216 polypeptide (e.g., the pre-, pro-, catalytic, cysteine-rich, disintegrin, EGF, transmembrane, and cytoplasmic domains) disclosed
5 herein.

Polypeptides according to the invention can comprise at least 5 amino acid residues; preferably the polypeptides comprise at least 12 residues; more preferably the polypeptides comprise at least 20 residues; and yet more
10 preferably the polypeptides comprise at least 30 residues. Nucleic acids comprising protein-coding sequences can be used to direct the expression of asthma-associated polypeptides in intact cells or in cell-free translation systems. The coding sequence can be tailored, if desired, for more efficient expression in a given host organism, and can be used to synthesize
15 oligonucleotides encoding the desired amino acid sequences. The resulting oligonucleotides can be inserted into an appropriate vector and expressed in a compatible host organism or translation system.

The polypeptides of the present invention, including function-conservative variants, may be isolated from wild-type or mutant cells (e.g.,
20 human cells or cell lines), from heterologous organisms or cells (e.g., bacteria, yeast, insect, plant, and mammalian cells), or from cell-free translation systems (e.g., wheat germ, microsomal membrane, or bacterial extracts) in which a protein-coding sequence has been introduced and expressed. Furthermore, the polypeptides may be part of recombinant fusion proteins. The polypeptides
25 can also, advantageously, be made by synthetic chemistry. Polypeptides may be chemically synthesized by commercially available automated procedures, including, without limitation, exclusive solid phase synthesis, partial solid phase methods, fragment condensation or classical solution synthesis.

Methods for polypeptide purification are well-known in the art, including,
30 without limitation, preparative disc-gel electrophoresis, isoelectric focusing, HPLC, reversed-phase HPLC, gel filtration, ion exchange and partition

chromatography, and countercurrent distribution. For some purposes, it is preferable to produce the polypeptide in a recombinant system in which the protein contains an additional sequence (e.g., epitope or protein) tag that facilitates purification. Non-limiting examples of epitope tags include c-myc, 5 haemagglutinin (HA), polyhistidine (6X-HIS) (SEQ ID NO:32), GLU-GLU, and DYKDDDDK (SEQ ID NO:33) (FLAG®) epitope tags. Non-limiting examples of protein tags include glutathione-S-transferase (GST), green fluorescent protein (GFP), and maltose binding protein (MBP).

In one approach, the coding sequence of a polypeptide or peptide can 10 be cloned into a vector that creates a fusion with a sequence tag of interest. Suitable vectors include, without limitation, pRSET (Invitrogen Corp., San Diego, CA), pGEX (Amersham-Pharmacia Biotech, Inc., Piscataway, NJ), pEGFP (CLONTECH Laboratories, Inc., Palo Alto, CA), and pMAL™ (New England BioLabs (NEB), Inc., Beverly, MA) plasmids. Following expression, 15 the epitope, or protein tagged polypeptide or peptide can be purified from a crude lysate of the translation system or host cell by chromatography on an appropriate solid-phase matrix. In some cases, it may be preferable to remove the epitope or protein tag (i.e., via protease cleavage) following purification. As an alternative approach, antibodies produced against a disorder-associated 20 protein or against peptides derived therefrom can be used as purification reagents. Other purification methods are possible.

The present invention also encompasses polypeptide derivatives of Gene 216. The isolated polypeptides may be modified by, for example, phosphorylation, sulfation, acylation, or other protein modifications. They may 25 also be modified with a label capable of providing a detectable signal, either directly or indirectly, including, but not limited to, radioisotopes and fluorescent compounds.

Both the naturally occurring and recombinant forms of the polypeptides of the invention can advantageously be used to screen compounds for binding 30 activity. Many methods of screening for binding activity are known by those skilled in the art and may be used to practice the invention. Several methods

of automated assays have been developed in recent years so as to permit screening of tens of thousands of compounds in a short period of time. Such high-throughput screening methods are particularly preferred. The use of high-throughput screening assays to test for inhibitors is greatly facilitated by the availability of large amounts of purified polypeptides, as provided by the invention. The polypeptides of the invention also find use as therapeutic agents as well as antigenic components to prepare antibodies.

The polypeptides of this invention find use as immunogenic components useful as antigens for preparing antibodies by standard methods. It is well known in the art that immunogenic epitopes generally contain at least about five amino acid residues (Ohno et al., 1985, *Proc. Natl. Acad. Sci. USA* 82:2945). Therefore, the immunogenic components of this invention will typically comprise at least 5 amino acid residues of the sequence of the complete polypeptide chains. Preferably, they will contain at least 7, and most preferably at least about 10 amino acid residues or more to ensure that they will be immunogenic. Whether a given component is immunogenic can readily be determined by routine experimentation. Such immunogenic components can be produced by proteolytic cleavage of larger polypeptides or by chemical synthesis or recombinant technology and are thus not limited by proteolytic cleavage sites. The present invention thus encompasses antibodies that specifically recognize asthma-associated immunogenic components.

Structural Studies

A purified Gene 216 polypeptide can be analyzed by well-established methods (e.g., X-ray crystallography, NMR, CD, etc.) to determine the three-dimensional structure of the molecule. The three-dimensional structure, in turn, can be used to model intermolecular interactions. Exemplary methods for crystallization and X-ray crystallography are found in P.G. Jones, 1981, *Chemistry in Britain*, 17:222-225; C. Jones et al. (eds), *Crystallographic Methods and Protocols*, Humana Press, Totowa, NJ; A. McPherson, 1982, *Preparation and Analysis of Protein Crystals*, John Wiley & Sons, New York, NY; T.L. Blundell and L.N. Johnson, 1976, *Protein Crystallography*, Academic

- Press, Inc., New York, NY; A. Holden and P. Singer, 1960, *Crystals and Crystal Growing*, Anchor Books-Doubleday, New York, NY; R.A. Laudise, 1970, *The Growth of Single Crystals*, Solid State Physical Electronics Series, N. Holonyak, Jr., (ed), Prentice-Hall, Inc.; G.H. Stout and L.H. Jensen, 1989, *X-ray Structure Determination: A Practical Guide*, 2nd edition, John Wiley & Sons, New York, NY; *Fundamentals of Analytical Chemistry*, 3rd. edition, Saunders Golden Sunburst Series, Holt, Rinehart and Winston, Philadelphia, PA, 1976; P.D. Boyle of the Department of Chemistry of North Carolina State University at <http://laue.chem.ncsu.edu/web/GrowXtal.html>; M.B. Berry, 1995, *Protein Crystalization: Theory and Practice, Structure and Dynamics of E. coli Adenylate Kinase*, Doctoral Thesis, Rice University, Houston TX; www.bioc.rice.edu/~berry/papers/crystalization/crystalization.html.

For X-ray diffraction studies, single crystals can be grown to suitable size. Preferably, a crystal has a size of 0.2 to 0.4 mm in at least two of the three dimensions. Crystals can be formed in a solution comprising a Gene 216 polypeptide (e.g., 1.5-200 mg/ml) and reagents that reduce the solubility to conditions close to spontaneous precipitation. Factors that affect the formation of polypeptide crystals include: 1) purity; 2) substrates or co-factors; 3) pH; 4) temperature; 5) polypeptide concentration; and 6) characteristics of the precipitant. Preferably, the Gene 216 polypeptides are pure, i.e., free from contaminating components (at least 95% pure), and free from denatured Gene 216 polypeptides. In particular, polypeptides can be purified by FPLC and HPLC techniques to assure homogeneity (see, Lin et al., 1992, *J. Crystal Growth*. 122:242-245). Optionally, Gene 216 polypeptide substrates or co-factors can be added to stabilize the quaternary structure of the protein and promote lattice packing.

Suitable precipitants for crystallization include, but are not limited to, salts (e.g., ammonium sulphate, potassium phosphate); polymers (e.g., polyethylene glycol (PEG) 6000); alcohols (e.g., ethanol); polyalcohols (e.g., 1-methyl-2,4 pentane diol (MPD)); organic solvents; sulfonic dyes; and deionized water. The ability of a salt to precipitate polypeptides can be

generally described by the Hofmeister series: $\text{PO}_4^{3-} > \text{HPO}_4^{2-} = \text{SO}_4^{2-} > \text{citrate} > \text{CH}_3\text{CO}_2^- > \text{Cl}^- > \text{Br}^- > \text{NO}_3^- > \text{ClO}_4^- > \text{SCN}^-$; and $\text{NH}_4^+ > \text{K}^+ > \text{Na}^+ > \text{Li}^+$. Non-limiting examples of salt precipitants are shown below (see Berry, 1995).

Precipitant	Maximum concentration
$(\text{NH}_4^+/\text{Na}^+/\text{Li}^+)_2$ or $\text{Mg}_2+\text{SO}_4^{2-}$	4.0 / 1.5 / 2.1 / 2.5 M
$\text{NH}_4^+/\text{Na}^+/\text{K}^+ \text{PO}_4^{3-}$	3.0 / 4.0 / 4.0 M
$\text{NH}_4^+/\text{K}^+/\text{Na}^+/\text{Li}^+ \text{citrate}$	~1.8 M
$\text{NH}_4^+/\text{K}^+/\text{Na}^+/\text{Li}^+ \text{acetate}$	~3.0 M
$\text{NH}_4^+/\text{K}^+/\text{Na}^+/\text{Li}^+ \text{Cl}^-$	5.2 / 9.8 / 4.2 / 5.4 M
$\text{NH}_4^+ \text{NO}_3^-$	~8.0 M

- 5 High molecular weight polymers useful as precipitating agents include polyethylene glycol (PEG), dextran, polyvinyl alcohol, and polyvinyl pyrrolidone (A. Polson et al., 1964, *Biochem. Biophys. Acta.* **82**:463-475). In general, polyethylene glycol (PEG) is the most effective for forming crystals. PEG compounds with molecular weights less than 1000 can be used at concentrations above 40% v/v. PEGs with molecular weights above 1000 can be used at concentration 5-50% w/v. Typically, PEG solutions are mixed with ~0.1 % sodium azide to prevent bacterial growth.

- Typically, crystallization requires the addition of buffers and a specific salt content to maintain the proper pH and ionic strength for a protein's stability.
- 15 Suitable additives include, but are not limited to sodium chloride (e.g., 50-500 mM as additive to PEG and MPD; 0.15-2 M as additive to PEG); potassium chloride (e.g., 0.05-2 M); lithium chloride (e.g., 0.05-2 M); sodium fluoride (e.g., 20-300 mM); ammonium sulfate (e.g., 20-300 mM); lithium sulfate (e.g., 0.05-2 M); sodium or ammonium thiocyanate (e.g., 50-500 mM); MPD (e.g., 0.5-50%);
- 20 1,6 hexane diol (e.g., 0.5-10%); 1,2,3 heptane triol (e.g., 0.5-15%); and benzamidine (e.g., 0.5-15%).

- Detergents may be used to maintain protein solubility and prevent aggregation. Suitable detergents include, but are not limited to non-ionic detergents such as sugar derivatives, oligoethyleneglycol derivatives,
- 25 dimethylamine-N-oxides, cholate derivatives, N-octyl hydroxyalkylsulphoxides, sulphobetains, and lipid-like detergents. Sugar-derived detergents include alkyl glucopyranosides (e.g., C8-GP, C9-GP), alkyl thio-glucopyranosides (e.g., C8-

tGP), alkyl maltopyranosides (e.g., C10-M, C12-M; CYMAL-3, CYMAL-5, CYMAL-6), alkyl thio-maltopyranosides, alkyl galactopyranosides, alkyl sucroses (e.g., N-octanoylsucrose), and glucamides (e.g., HECAMEG, C-HEGA-10; MEGA-8). Oligoethyleneglycol-derived detergents include alkyl
5 polyoxyethylenes (e.g., C8-E5, C8-En; C12-E8; C12-E9) and phenyl polyoxyethylenes (e.g., Triton X-100). Dimethylamine-N-oxide detergents include, e.g., C10-DAO; DDAO; LDAO. Cholate-derived detergents include, e.g., Deoxy-Big CHAP, digitonin. Lipid-like detergents include phosphocholine compounds. Suitable detergents further include zwitter-ionic detergents (e.g.,
10 ZWITTERGENT 3-10; ZWITTERGENT 3-12); and ionic detergents (e.g., SDS).

Crystallization of macromolecules has been performed at temperatures ranging from 60°C to less than 0°C. However, most molecules can be crystallized at 4°C or 22°C. Lower temperatures promote stabilization of polypeptides and inhibit bacterial growth. In general, polypeptides are more
15 soluble in salt solutions at lower temperatures (e.g., 4°C), but less soluble in PEG and MPD solutions at lower temperatures. To allow crystallization at 4°C or 22°C, the precipitant or protein concentration can be increased or decreased as required. Heating, melting, and cooling of crystals or aggregates can be used to enlarge crystals. In addition, crystallization at both 4°C and 22°C can
20 be assessed (A. McPherson, 1992, *J. Cryst. Growth*. **122**:161-167; C.W. Carter, Jr. and C.W. Carter, 1979, *J. Biol. Chem.* **254**:12219-12223; T. Bergfors, 1993, *Crystallization Lab Manual*).

A crystallization protocol can be adapted to a particular polypeptide or peptide. In particular, the physical and chemical properties of the polypeptide
25 can be considered (e.g., aggregation, stability, adherence to membranes or tubing, internal disulfide linkages, surface cysteines, chelating ions, etc.). For initial experiments, the standard set of crystallization reagents can be used (Hampton Research, Laguna Niguel, CA). In addition, the CRYSTOOL program can provide guidance in determining optimal crystallization conditions
30 (Brent Segelke, 1995, Efficiency analysis of sampling protocols used in protein crystallization screening and crystal structure from two novel crystal forms of

PLA2, Ph.D. Thesis, University of California, San Diego; <http://www.ccp14.ac.uk/ccp/web-mirrors/lnlrupp/crystool/crystool.htm>). Exemplary crystallization conditions are shown below (see Berry, 1995).

Major Precipitant	Additive	Concentration of Major Precipitant	Concentration of Additive
(NH ₄) ₂ SO ₄	PEG 400-2000, MPD, ethanol, or methanol	2.0-4.0 M	6%-0.5%
Na citrate	PEG 400-2000, MPD, ethanol, or methanol	1.4-1.8 M	6%-0.5%
PEG 1000-20000	(NH ₄) ₂ SO ₄ , NaCl, or Na formate	40-50%	0.2-0.6 M

5

Robots can be used for automatic screening and optimization of crystallization conditions. For example, the IMPAX and Oryx systems can be used (Douglas Instruments, Ltd., East Garston, United Kingdom). The CRYSTOOL program (Segelke, *supra*) can be integrated with the robotics programming. In addition, the Xact program can be used to construct, maintain, and record the results of various crystallization experiments (see, e.g., D.E. Brodersen et al., 1999, *J. Appl. Cryst.* **32**: 1012-1016; G.R. Andersen and J. Nyborg, 1996, *J. Appl. Cryst.* **29**:236-240). The Xact program supports multiple users and organizes the results of crystallization experiments into hierarchies. Advantageously, Xact is compatible with both CRYSTOOL and Microsoft® Excel programs.

Four methods are commonly employed to crystallize macromolecules: vapor diffusion, free interface diffusion, batch, and dialysis. The vapor diffusion technique is typically performed by formulating a 1:1 mixture of a solution comprising the polypeptide of interest and a solution containing the precipitant at the final concentration that is to be achieved after vapor equilibration. The drop containing the 1:1 mixture of protein and precipitant is then suspended and sealed over the well solution, which contains the precipitant at the target concentration, as either a hanging or sitting drop. Vapor diffusion can be used to screen a large number of crystallization conditions or when small amounts of polypeptide are available. For screening, drop sizes of 1 to 2 μ l can be used. Once preliminary crystallization conditions

have been determined, drop sizes such as 10 μ l can be used. Notably, results from hanging drops may be improved with agarose gels (see K. Provost and M.-C. Robert, 1991, *J. Cryst. Growth*. **110**:258-264). Free interface diffusion is performed by layering of a low density solution onto one of higher density, usually in the form of concentrated protein onto concentrated salt. Since the solute to be crystallized must be concentrated, this method typically requires relatively large amounts of protein. However, the method can be adapted to work with small amounts of protein. In a representative experiment, 2 to 5 μ l of sample is pipetted into one end of a 20 μ l microcapillary pipet. Next, 2 to 5 μ l of precipitant is pipetted into the capillary without introducing an air bubble, and the ends of the pipet are sealed. With sufficient amounts of protein, this method can be used to obtain relatively large crystals (see, e.g., S.M. Althoff et al., 1988, *J. Mol. Biol.* **199**:665-666).

The batch technique is performed by mixing concentrated polypeptide with concentrated precipitant to produce a final concentration that is supersaturated for the solute macromolecule. Notably, this method can employ relatively large amounts of solution (e.g., milliliter quantities), and can produce large crystals. For that reason, the batch technique is not recommended for screening initial crystallization conditions.

The dialysis technique is performed by diffusing precipitant molecules through a semipermeable membrane to slowly increase the concentration of the solute inside the membrane. Dialysis tubing can be used to dialyze milliliter quantities of sample, whereas dialysis buttons can be used to dialyze microliter quantities (e.g., 7-200 μ l). Dialysis buttons may be constructed out of glass, perspex, or Teflon™ (see, e.g., Cambridge Repetition Engineers Ltd., Greens Road, Cambridge CB4 3EQ, UK; Hampton Research). Using this method, the precipitating solution can be varied by moving the entire dialysis button or sack into a different solution. In this way, polypeptides can be "reused" until the correct conditions for crystallization are found (see, e.g., C.W. Carter, Jr. et al., 1988, *J. Cryst. Growth*. **90**:60-73). However, this method is not recommended for precipitants comprising concentrated PEG solutions.

Various strategies have been designed to screen crystallization conditions, including 1) pl screening; 2) grid screening; 3) factorials; 4) solubility assays; 5) perturbation; and 6) sparse matrices. In accordance with the pl screening method, the pl of a polypeptide is presumed to be its
5 crystallization point. Screening at the pl can be performed by dialysis against low concentrations of buffer (less than 20 mM) at the appropriate pH, or by use of conventional precipitants.

The grid screening method can be performed on two-dimensional matrices. Typically, the precipitant concentration is plotted against pH. The
10 optimal conditions can be determined for each axis, and then combined. At that point, additional factors can be tested (e.g., temperature, additives). This method works best with fast-forming crystals, and can be readily automated (see M.J. Cox and P.C. Weber, 1988, *J. Cryst. Growth*. **90**:318-324). Grid screens are commercially available for popular precipitants such as ammonium
15 sulphate, PEG 6000, MPD, PEG/LiCl, and NaCl (see, e.g., Hamilton Research).

The incomplete factorial method can be performed by 1) selecting a set of ~20 conditions; 2) randomly assigning combinations of these conditions; 3) grading the success of the results of each experiment using an objective scale;
20 and 4) statistically evaluating the effects of each of the conditions on crystal formation (see, e.g., C.W. Carter, Jr. et al., 1988, *J. Cryst. Growth*. **90**:60-73).

In particular, conditions such as pH, temperature, precipitating agent, and cations can be tested. Dialysis buttons are preferably used with this method. Typically, optimal conditions/combinations can be determined within 35 tests.
25 Similar approaches, such as "footprinting" conditions, may also be employed (see, e.g., E.A. Stura et al., 1991, *J. Cryst. Growth*. **110**:1-2).

The perturbation approach can be performed by altering crystallization conditions by introducing a series of additives designed to test the effects of altering the structure of bulk solvent and the solvent dielectric on crystal
30 formation (see, e.g., Whitaker et al., 1995, *Biochem*. **34**:8221-8226). Additives for increasing the solvent dielectric include, but are not limited to, NaCl, KCl,

or LiCl (e.g., 200 mM); Na formate (e.g., 200 mM); Na_2HPO_4 or K_2HPO_4 (e.g., 200 mM); urea, trichloroacetate, guanidium HCl, or KSCN (e.g., 20-50 mM).

- A non-limiting list of additives for decreasing the solvent dielectric include methanol, ethanol, isopropanol, or tert-butanol (e.g., 1-5%); MPD (e.g., 1%);
 5 PEG 400, PEG 600, or PEG 1000 (e.g., 1-4%); PEG MME (monomethylether) 550, PEG MME 750, PEG MME 2000 (e.g., 1-4%).

- As an alternative to the above-screening methods, the sparse matrix approach can be used (see, e.g., J. Jancarik and S.-H.J. Kim, 1991, *Appl. Cryst.* **24**:409-411; A. McPherson, 1992, *J. Cryst. Growth.* **122**:161-167; B.
 10 Cudney et al., 1994, *Acta. Cryst.* **D50**:414-423). Sparse matrix screens are commercially available (see, e.g., Hampton Research; Molecular Dimensions, Inc., Apopka, FL; Emerald Biostructures, Inc., Lemont, IL). Notably, data from Hampton Research sparse matrix screens can be stored and analyzed using ASPRUN software (Douglas Instruments).

- 15 Exemplary conditions for an initial screen are shown below (see Berry, 1995).

TABLE 1

Tray 1:

PEG 8000 (wells 1-6)						Ammonium sulfate (wells 7-12)					
1	2	3	4	5	6	7	8	9	10	11	12
20% pH 5.0	20% pH 7.0	20% pH 8.8	35% pH 5.0	35% pH 7.0	35% pH 8.8	2.0 M pH 5.0	2.0 M pH 7.0	2.0 M pH 8.8	2.5 M pH 5.0	2.5 M pH 7.0	2.5 M pH 8.8
MPD (wells 13-16)				Na Citrate (wells 17-20)				Na/K Phosphate (wells 21-24)			
13	14	15	16	17	18	19	20	21	22	23	24
30% pH 5.8	30% pH 7.6	50% pH 5.8	50% pH 7.6	1.3 M pH 6.8	1.3 M pH 7.5	1.5 M pH 5.8	1.5 M pH 7.5	2.0 M pH 6.0	2.0 M pH 7.4	2.5 M pH 6.0	2.5 M pH 7.4

20 Tray 2:

PEG 2000 MME/0.2 M Ammon. sulfate (wells 25-30)					
25	26	27	28	29	30
25% pH 5.5	25% pH 7.0	25% pH 8.5	40% pH 5.5	40% pH 7.0	40% pH 8.5
Random for wells 31 to 48					

- The initial screen can be used with hanging or sitting drops. To conserve the sample, tray 2 can be set up several weeks following tray 1. Wells 31-48 of tray 2 can comprise a random set of solutions. Alternatively,
 25 solutions can be formulated using sparse methods. Preferably, test solutions cover a broad range of precipitants, additives, and pH (especially pH 5.0-9.0).

Seeding can be used to trigger nucleation and crystal growth (Stura and

Wilson, 1990, *J. Cryst. Growth*. **110**:270-282; C. Thaller et al., 1981, *J. Mol. Biol.* **147**:465-469; A. McPherson and P. Schlichta, 1988, *J. Cryst. Growth*. **90**:47-50). In general, seeding can be performed by transferring crystal seeds into a polypeptide solution to allow polypeptide molecules to deposit on the surface of the seeds and produce crystals. Two seeding methods can be used: microseeding and macroseeding. For microseeding, a crystal can be ground into tiny pieces and transferred into the protein solution. Alternatively, seeds can be transferred by adding 1-2 μ l of the seed solution directly to the equilibrated protein solution. In another approach, seeds can be transferred by dipping a hair in the seed solution and then streaking the hair across the surface of the drop (streak seeding; see Stura and Wilson, *supra*). For macroseeding, an intact crystal can be transferred into the protein solution (see, e.g., C. Thaller et al., 1981, *J. Mol. Biol.* **147**:465-469). Preferably, the surface of the crystal seed is washed to regenerate the growing surface prior to being transferred. Optimally, the protein solution for crystallization is close to saturation and the crystal seed is not completely dissolved upon transfer.

Antibodies

An isolated Gene 216 polypeptide or a portion or fragment thereof, can be used as an immunogen to generate anti-Gene 216 antibodies using standard techniques for polyclonal and monoclonal antibody preparation. The full-length Gene 216 polypeptide can be used or, alternatively, the invention provides antigenic peptide fragments of Gene 216 for use as immunogens. The antigenic peptide of Gene 216 comprises at least 5 amino acid residues of the amino acid sequence shown in SEQ ID NO:4, and encompasses an epitope of Gene 216 such that an antibody raised against the peptide forms a specific immune complex with Gene 216 amino acid sequence.

Accordingly, another aspect of the invention pertains to anti-Gene 216 antibodies. The invention provides polyclonal and monoclonal antibodies that bind Gene 216 polypeptides or peptides. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site

capable of immunoreacting with a particular epitope of a Gene 216 polypeptide or peptide. A monoclonal antibody composition thus typically displays a single binding affinity for a particular Gene 216 polypeptide or peptide with which it immunoreacts.

- 5 A Gene 216 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (e.g., rabbit, goat, mouse, or other non-human mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed Gene 216 polypeptide or a chemically synthesized Gene 216 polypeptide, or fragments thereof. The
10 preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic Gene 216 preparation induces a polyclonal anti-Gene 216 antibody response.

- A number of adjuvants are known and used by those skilled in the art.
- 15 Non-limiting examples of suitable adjuvants include incomplete Freund's adjuvant, mineral gels such as alum, aluminum phosphate, aluminum hydroxide, aluminum silica, and surface-active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Further examples of adjuvants include N-
20 acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3
hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and RIBI, which contains three components extracted from bacteria,
25 monophosphoryl lipid A, trehalose dimycolate and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 80 emulsion. A particularly useful adjuvant comprises 5% (wt/vol) squalene, 2.5% Pluronic L121 polymer and 0.2% polysorbate in phosphate buffered saline (Kwak et al., 1992, *New Eng. J. Med.* **327**:1209-1215). Preferred adjuvants include complete BCG, Detox,
30 (RIBI, Immunochem Research Inc.), ISCOMS, and aluminum hydroxide adjuvant (Superphos, Biosector). The effectiveness of an adjuvant may be

determined by measuring the amount of antibodies directed against the immunogenic peptide.

Polyclonal anti-Gene 216 antibodies can be prepared as described above by immunizing a suitable subject with a Gene 216 immunogen. The
5 anti-Gene 216 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized Gene 216. If desired, the antibody molecules directed against Gene 216 can be isolated from the mammal (e.g., from the blood) and further purified by well-known techniques, such as protein A
10 chromatography to obtain the IgG fraction.

At an appropriate time after immunization, e.g., when the anti-Gene 216 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique (see Kohler and Milstein, 1975, *Nature*
15 **256**:495-497; Brown et al., 1981, *J. Immunol.* **127**:539-46; Brown et al., 1980, *J. Biol. Chem.* **255**:4980-83; Yeh et al., 1976, *PNAS* **76**:2927-31; and Yeh et al., 1982, *Int. J. Cancer* **29**:269-75), the human B cell hybridoma technique (Kozbor et al., 1983, *Immunol. Today* **4**:72), the EBV-hybridoma technique (Cole et al., 1985, *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss,
20 Inc., pp. 77-96) or trioma techniques.

The technology for producing hybridomas is well-known (see generally R. H. Kenneth, 1980, *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, NY; E.A. Lerner, 1981, *Yale J. Biol. Med.*, **54**:387-402; M.L. Geffer et al., 1977, *Somatic Cell Genet.* **3**:231-
25 36). In general, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a Gene 216 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds Gene 216 polypeptides or peptides.

30 Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-

Gene 216 monoclonal antibody (see, e.g., G. Galfre et al., 1977, *Nature* 266:55052; Gefter et al., 1977; Lerner, 1981; Kenneth, 1980). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods. Typically, the immortal cell line (e.g., a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin, and thymidine (HAT medium). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, e.g., the P3-NS1/1-Ag4-1, P3-x63-Ag8.653, or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC (American Type Culture Collection, Manassas, VA). Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol (PEG). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind Gene 216 polypeptides or peptides, e.g., using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-Gene 216 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with Gene 216 to thereby isolate immunoglobulin library members that bind Gene 216. Kits for generating and screening phage display libraries are commercially available (e.g., the Pharmacia Recombinant Phage Antibody System, Catalog No. 27-9400-01; and the Stratagene SurfZAP™ Phage Display Kit, Catalog No. 240612).

Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in,

- for example, Ladner et al. U.S. Pat. No. 5,223,409; Kang et al. PCT International Publication No. WO 92/18619; Dower et al. PCT International Publication No. WO 91/17271; Winter et al. PCT International Publication WO 92/20791; Markland et al. PCT International Publication No. WO 92/15679;
- 5 Breittling et al. PCT International Publication WO 93/01288; McCafferty et al. PCT International Publication No. WO 92/01047; Garrard et al. PCT International Publication No. WO 92/09690; Ladner et al. PCT International Publication No. WO 90/02809; Fuchs et al., 1991, *Bio/Technology* 9:1370-1372; Hay et al., 1992, *Hum. Antibod. Hybridomas* 3:81-85; Huse et al., 1989,
- 10 *Science* 246:1275-1281; Griffiths et al., 1993, *EMBO J* 12:725-734; Hawkins et al., 1992, *J. Mol. Biol.* 226:889-896; Clarkson et al., 1991, *Nature* 352:624-628; Gram et al., 1992, *PNAS* 89:3576-3580; Garrad et al., 1991, *Bio/Technology* 9:1373-1377; Hoogenboom et al., 1991, *Nuc. Acid Res.* 19:4133-4137; Barbas et al., 1991, *PNAS* 88:7978-7982; and McCafferty et al.,
- 15 1990, *Nature* 348:552-55.

Additionally, recombinant anti-Gene 216 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and

20 humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson et al. International Application No. PCT/US86/02269; Akira, et al. European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al. European Patent Application 173,494; Neuberger et

25 al. PCT International Publication No. WO 86/01533; Cabilly et al. U.S. Pat. No. 4,816,567; Cabilly et al. European Patent Application 125,023; Better et al., 1988, *Science* 240:1041-1043; Liu et al., 1987, *PNAS* 84:3439-3443; Liu et al., 1987, *J. Immunol.* 139:3521-3526; Sun et al., 1987, *PNAS* 84:214-218; Nishimura et al., 1987, *Canc. Res.* 47:999-1005; Wood et al., 1985, *Nature*

30 314:446-449; and Shaw et al., 1988, *J. Natl. Cancer Inst.* 80:1553-1559; S.L. Morrison, 1985, *Science* 229:1202-1207; Oi et al., 1986, *BioTechniques* 4:214;

Winter U.S. Pat. No. 5,225,539; Jones et al., 1986, *Nature* **321**:552-525; Verhoeyan et al., 1988, *Science* **239**:1534; and Bcidler et al., 1988, *J. Immunol.* **141**:4053-4060.

An anti-Gene 216 antibody (e.g., monoclonal antibody) can be used to
5 isolate Gene 216 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-Gene 216 antibody can also facilitate the purification of natural Gene 216 polypeptide from cells and of recombinantly produced Gene 216 polypeptides or peptides expressed in host cells. Further, an anti-Gene 216 antibody can be used to detect Gene 216 protein (e.g., in a
10 cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the Gene 216 protein. Anti-Gene 216 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen as described in detail herein. In addition, and anti-Gene
15 216 antibody can be used as therapeutics for the treatment of diseases related to abnormal Gene 216 expression or function, e.g., asthma.

Ligands

The Gene 216 polypeptides, polynucleotides, variants, or fragments thereof, can be used to screen for ligands (e.g., agonists, antagonists, or
20 inhibitors) that modulate the levels or activity of the Gene 216 polypeptide. In addition, these Gene 216 molecules can be used to identify endogenous ligands that bind to Gene 216 polypeptides or polynucleotides in the cell. In one aspect of the present invention, the full-length Gene 216 polypeptide (e.g., SEQ ID NO:4) is used to identify ligands. Alternatively, variants or fragments
25 of a Gene 216 polypeptide are used. Such fragments may comprise, for example, one or more domains of the Gene 216 polypeptide (e.g., the pre-, pro-, catalytic, cysteine-rich, disintegrin, EGF, transmembrane, and cytoplasmic domains) disclosed herein. Of particular interest are screening assays that identify agents that have relatively low levels of toxicity in human cells. A wide
30 variety of assays may be used for this purpose, including *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays, and the

like.

The term "ligand" as used herein describes any molecule, protein, peptide, or compound with the capability of directly or indirectly altering the physiological function, stability, or levels of the Gene 216 polypeptide. Ligands that bind to the Gene 216 polypeptides or polynucleotides of the invention are potentially useful in diagnostic applications and/or pharmaceutical compositions, as described in detail herein. Ligands may encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 daltons. Such ligands can comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. Ligands often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Ligands can also comprise biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs, or combinations thereof.

Ligands may include, for example, 1) peptides such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, e.g., Lam et al., 1991, *Nature* **354**:82-84; Houghten et al., 1991, *Nature* **354**:84-86) and combinatorial chemistry-derived molecular libraries made of D- and/or L-configuration amino acids; 2) phosphopeptides (e.g., members of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang et al, 1993, *Cell* **72**:767-778); 3) antibodies (e.g., polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, F(ab')₂, Fab expression library fragments, and epitope-binding fragments of antibodies); and 4) small organic and inorganic molecules.

Ligands can be obtained from a wide variety of sources including libraries of synthetic or natural compounds. Synthetic compound libraries are commercially available from, for example, Maybridge Chemical Co. (Trevillet,

Cornwall, UK), Comgenex (Princeton, NJ), Brandon Associates (Merrimack, NH), and Microsource (New Milford, CT). A rare chemical library is available from Aldrich Chemical Company, Inc. (Milwaukee, WI). Natural compound libraries comprising bacterial, fungal, plant or animal extracts are available
5 from, for example, Pan Laboratories (Bothell, WA). In addition, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides.

Alternatively, libraries of natural compounds in the form of bacterial,
10 fungal, plant and animal extracts can be readily produced. Methods for the synthesis of molecular libraries are readily available (see, e.g., DeWitt et al., 1993, *Proc. Natl. Acad. Sci. USA* **90**:6909; Erb et al., 1994, *Proc. Natl. Acad. Sci. USA* **91**:11422; Zuckermann et al., 1994, *J. Med. Chem.* **37**:2678; Cho et al., 1993, *Science* **261**:1303; Carell et al., 1994, *Angew. Chem. Int. Ed. Engl.*
15 **33**:2059; Carell et al., 1994, *Angew. Chem. Int. Ed. Engl.* **33**:2061; and in Gallop et al., 1994, *J. Med. Chem.* **37**:1233). In addition, natural or synthetic compound libraries and compounds can be readily modified through conventional chemical, physical and biochemical means (see, e.g., Blondelle et al., 1996, *Trends in Biotech.* **14**:60), and may be used to produce
20 combinatorial libraries. In another approach, previously identified pharmacological agents can be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, and the analogs can be screened for Gene 216-modulating activity.

Numerous methods for producing combinatorial libraries are known in
25 the art, including those involving biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four
30 approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds (K. S. Lam, 1997, *Anticancer Drug Des.*

12:145).

Libraries may be screened in solution (e.g., Houghten, 1992, *Biotechniques* **13**:412-421), or on beads (Lam, 1991, *Nature* **354**:82-84), chips (Fodor, 1993, *Nature* **364**:555-556), bacteria or spores (Ladner U.S. Pat. No. 5,223,409), plasmids (Cull et al., 1992, *Proc. Natl. Acad. Sci. USA* **89**:1865-1869), or on phage (Scott and Smith, 1990, *Science* **249**:386-390; Devlin, 1990, *Science* **249**:404-406; Cwirla et al., 1990, *Proc. Natl. Acad. Sci. USA* **97**:6378-6382; Felici, 1991, *J. Mol. Biol.* **222**:301-310; Ladner, *supra*).

Where the screening assay is a binding assay, a Gene 216 polypeptide, polynucleotide, analog, or fragment thereof, may be joined to a label, where the label can directly or indirectly provide a detectable signal. Various labels include radioisotopes, fluorescers, chemilumescers, enzymes, specific binding molecules, particles, e.g. magnetic particles, and the like. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin, etc. For the specific binding members, the complementary member would normally be labeled with a molecule that provides for detection, in accordance with known procedures.

A variety of other reagents may be included in the screening assay. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc., that are used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Reagents that improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The components are added in any order that produces the requisite binding. Incubations are performed at any temperature that facilitates optimal activity, typically between 4° and 40°C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high-throughput screening. Normally, between 0.1 and 1 hr will be sufficient. In general, a plurality of assay mixtures is run in parallel with different agent concentrations to obtain a differential response to these concentrations. Typically, one of these concentrations serves as a negative control, i.e. at zero concentration or below the level of detection.

To perform cell-free ligand screening assays, it may be desirable to immobilize either the Gene 216 polypeptide, polynucleotide, or fragment to a surface to facilitate identification of ligands that bind to these molecules, as well as to accommodate automation of the assay. For example, a fusion protein comprising a Gene 216 polypeptide and an affinity tag can be produced. In one embodiment, a glutathione-S-transferase/phosphodiesterase fusion protein comprising a Gene 216 polypeptide is adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione-derivatized microtiter plates. Cell lysates (e.g., containing ³⁵S-labeled polypeptides) are added to the Gene 216-coated beads under conditions to allow complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the Gene 216-coated beads are washed to remove any unbound polypeptides, and the amount of immobilized radiolabel is determined. Alternatively, the complex is dissociated and the radiolabel present in the supernatant is determined. In another approach, the beads are analyzed by SDS-PAGE to identify Gene 216-binding polypeptides.

Ligand-binding assays can be used to identify agonist or antagonists that alter the function or levels of the Gene 216 polypeptide. Such assays are designed to detect the interaction of test agents with Gene 216 polypeptides, polynucleotides, analogs, or fragments thereof. Interactions may be detected by direct measurement of binding. Alternatively, interactions may be detected by indirect indicators of binding, such as stabilization/destabilization of protein structure, or activation/inhibition of biological function. Non-limiting examples of useful ligand-binding assays are detailed below.

Ligands that bind to Gene 216 polypeptides, polynucleotides, analogs, or fragments thereof, can be identified using real-time Bimolecular Interaction Analysis (BIA; Sjolander et al., 1991, *Anal. Chem.* **63**:2338-2345; Szabo et al., 1995, *Curr. Opin. Struct. Biol.* **5**:699-705). BIA-based technology (e.g., BIAcore™, LKB Pharmacia, Sweden) allows study of biospecific interactions in real time, without labeling. In BIA, changes in the optical phenomenon surface plasmon resonance (SPR) is used determine real-time interactions of

biological molecules.

Ligands can also be identified by scintillation proximity assays (SPA, described in U.S. Patent No. 4,568,649). In a modification of this assay that is currently undergoing development, chaperonins are used to distinguish
5 folded and unfolded proteins. A tagged protein is attached to SPA beads, and test agents are added. The bead is then subjected to mild denaturing conditions (such as, e.g., heat, exposure to SDS, etc.) and a purified labeled chaperonin is added. If a test agent binds to a target, the labeled chaperonin will not bind; conversely, if no test agent binds, the protein will undergo some
10 degree of denaturation and the chaperonin will bind.

Ligands can also be identified using a binding assay based on mitochondrial targeting signals (Hurt et al., 1985, *EMBO J.* 4:2061-2068; Eilers and Schatz, 1986, *Nature* 322:228-231). In a mitochondrial import assay, expression vectors are constructed in which nucleic acids encoding particular
15 target proteins are inserted downstream of sequences encoding mitochondrial import signals. The chimeric proteins are synthesized and tested for their ability to be imported into isolated mitochondria in the absence and presence of test compounds. A test compound that binds to the target protein should inhibit its uptake into isolated mitochondria *in vitro*.

20 The ligand-binding assay described in Fodor et al., 1991, *Science* 251:767-773, which involves testing the binding affinity of test compounds for a plurality of defined polymers synthesized on a solid substrate, can also be used.

Ligands that bind to Gene 216 polypeptides or peptides can be
25 identified using two-hybrid assays (see, e.g., U.S. Pat. No. 5,283,317; Zervos et al., 1993, *Cell* 72:223-232; Madura et al., 1993, *J. Biol. Chem.* 268:12046-12054; Bartel et al., 1993, *Biotechniques* 14:920-924; Iwabuchi et al., 1993, *Oncogene* 8:1693-1696; and Brent WO 94/10300). The two-hybrid system relies on the reconstitution of transcription activation activity by association of
30 the DNA-binding and transcription activation domains of a transcriptional activator through protein-protein interaction. The yeast GAL4 transcriptional

activator may be used in this way, although other transcription factors have been used and are well known in the art. To carryout the two-hybrid assay, the GAL4 DNA-binding domain, and the GAL4 transcription activation domain are expressed, separately, as fusions to potential interacting polypeptides.

- 5 In one embodiment, the "bait" protein comprises a Gene 216 polypeptide fused to the GAL4 DNA-binding domain. The "fish" protein comprises, for example, a human cDNA library encoded polypeptide fused to the GAL4 transcription activation domain. If the two, coexpressed fusion proteins interact in the nucleus of a host cell, a reporter gene (e.g. LacZ) is
- 10 activated to produce a detectable phenotype. The host cells that show two-hybrid interactions can be used to isolate the containing plasmids containing the cDNA library sequences. These plasmids can be analyzed to determine the nucleic acid sequence and predicted polypeptide sequence of the candidate ligand. Alternatively, methods such as the three-hybrid (Licitra et al.,
- 15 1996, *Proc. Natl. Acad. Sci. USA* **93**:12817-12821), and reverse two-hybrid (Vidal et al., 1996, *Proc. Natl. Acad. Sci. USA* **93**:10315-10320) systems may be used. Commercially available two-hybrid systems such as the CLONTECH Matchmaker™ systems and protocols (CLONTECH Laboratories, Inc., Palo Alto, CA) may be also be used (see also, A.R. Mendelsohn et al., 1994, *Curr.*
- 20 *Op. Biotech.* **5**:482; E.M. Phizicky et al., 1995, *Microbiological Rev.* **59**:94; M. Yang et al., 1995, *Nucleic Acids Res.* **23**:1152; S. Fields et al., 1994, *Trends Genet.* **10**:286; and U.S. Patent No. 6,283,173 and 5,468,614).

- Several methods of automated assays have been developed in recent years so as to permit screening of tens of thousands of test agents in a short
- 25 period of time. High-throughput screening methods are particularly preferred for use with the present invention. The ligand-binding assays described herein can be adapted for high-throughput screens, or alternative screens may be employed. For example, continuous format high throughput screens (CF-HTS) using at least one porous matrix allows the researcher to test large numbers
- 30 of test agents for a wide range of biological or biochemical activity (see United States Patent No. 5,976,813 to Beutel et al.). Moreover, CF-HTS can be used

to perform multi-step assays.

Diagnostics

As discussed herein, chromosomal region 20p13-p12 has been genetically linked to a variety of diseases and disorders, including asthma.

- 5 The present invention provides nucleic acids and antibodies that can be useful in diagnosing individuals with aberrant Gene 216 expression. In particular, the disclosed SNPs can be used to diagnose chromosomal abnormalities linked to these diseases.

- Antibody-based diagnostic methods: In a further embodiment of the present invention, antibodies which specifically bind to the Gene 216 polypeptide may be used for the diagnosis of conditions or diseases characterized by underexpression or overexpression of the Gene 216 polynucleotide or polypeptide, or in assays to monitor patients being treated with a Gene 216 polypeptide or peptide, or a Gene 216 agonist, antagonist, 15 or inhibitor.

The antibodies useful for diagnostic purposes may be prepared in the same manner as those for use in therapeutic methods, described herein.

- Antibodies may be raised to the full-length Gene 216 polypeptide sequence (e.g., SEQ ID NO:4). Alternatively, the antibodies may be raised to 20 fragments or variants of the Gene 216 polypeptide. In one aspect of the invention, antibodies are prepared to bind to a Gene 216 polypeptide fragment comprising one or more domains of the Gene 216 polypeptide (e.g., pre-, pro-, catalytic, disintegrin, cysteine-rich, EGF, transmembrane, and cytoplasmic domains) described herein.

- 25 Diagnostic assays for the Gene 216 polypeptide include methods that utilize the antibody and a label to detect the protein in biological samples (e.g., human body fluids, cells, tissues, or extracts of cells or tissues). The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule.

- 30 A wide variety of reporter molecules that are known in the art may be used, several of which are described herein.

The invention provides methods for detecting disease-associated antigenic components in a biological sample, which methods comprise the steps of: 1) contacting a sample suspected to contain a disease-associated antigenic component with an antibody specific for an disease-associated antigen, extracellular or intracellular, under conditions in which an antigen-antibody complex can form between the antibody and disease-associated antigenic components in the sample; and 2) detecting any antigen-antibody complex formed in step (1) using any suitable means known in the art, wherein the detection of a complex indicates the presence of disease-associated antigenic components in the sample. It will be understood that assays that utilize antibodies directed against altered Gene 216 amino acid sequences (i.e., epitopes encoded by SNPs, mutations, or variants) are within the scope of the invention.

Many immunoassay formats are known in the art, and the particular format used is determined by the desired application. An immunoassay can use, for example, a monoclonal antibody directed against a single disease-associated epitope, a combination of monoclonal antibodies directed against different epitopes of a single disease-associated antigenic component, monoclonal antibodies directed towards epitopes of different disease-associated antigens, polyclonal antibodies directed towards the same disease-associated antigen, or polyclonal antibodies directed towards different disease-associated antigens. Protocols can also, for example, use solid supports, or may involve immunoprecipitation.

In accordance with the present invention, "competitive" (U.S. Pat. Nos. 3,654,090 and 3,850,752), "sandwich" (U.S. Pat. No. 4,016,043), and "double antibody," or "DASP" assays may be used. Several procedures for measuring the Gene 216 polypeptide (e.g., ELISA, RIA, and FACS) are known in the art and provide a basis for diagnosing altered or abnormal levels of Gene 216 polypeptide expression. Normal or standard values for Gene 216 polypeptide expression are established by incubating biological samples taken from normal subjects, preferably human, with antibody to the Gene polypeptide under

conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods; photometric means are preferred. Levels of the Gene 216 polypeptide expressed in the subject sample, negative control (normal) sample, and positive control (disease) sample are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

Typically, immunoassays use either a labeled antibody or a labeled antigenic component (e.g., that competes with the antigen in the sample for binding to the antibody). A number of fluorescent materials are known and can be utilized as labels for antibodies or polypeptides. These include, for example, Cy3, Cy5, Alexa, BODIPY, fluorescein (e.g., FluorX, DTAF, and FITC), rhodamine (e.g., TRITC), auramine, Texas Red, AMCA blue, and Lucifer Yellow. Antibodies or polypeptides can also be labeled with a radioactive element or with an enzyme. Preferred isotopes include ^3H , ^{14}C , ^{32}P , ^{35}S , ^{36}Cl , ^{51}Cr , ^{57}Co , ^{58}Co , ^{59}Fe , ^{90}Y , ^{125}I , ^{131}I , and ^{186}Re . Preferred enzymes include peroxidase, β -glucuronidase, β -D-glucosidase, β -D-galactosidase, urease, glucose oxidase plus peroxidase, and alkaline phosphatase (see, e.g., U.S. Pat. Nos. 3,654,090; 3,850,752 and 4,016,043). Enzymes can be conjugated by reaction with bridging molecules such as carbodiimides, diisocyanates, glutaraldehyde, and the like. Enzyme labels can be detected visually, or measured by calorimetric, spectrophotometric, fluorospectrophotometric, amperometric, or gasometric techniques. Other labeling systems, such as avidin/biotin, Tyramide Signal Amplification (TSATM), are known in the art, and are commercially available (see, e.g., ABC kit, Vector Laboratories, Inc., Burlingame, CA; NEN[®] Life Science Products, Inc., Boston, MA).

Kits suitable for antibody-based diagnostic applications typically include one or more of the following components:

(1) Antibodies: The antibodies may be pre-labeled; alternatively, the antibody may be unlabeled and the ingredients for labeling may be included in the kit in separate containers, or a secondary, labeled antibody is

provided; and

(2) Reaction components: The kit may also contain other suitably packaged reagents and materials needed for the particular immunoassay protocol, including solid-phase matrices, if applicable, and standards.

- 5 The kits referred to above may include instructions for conducting the test. Furthermore, in preferred embodiments, the diagnostic kits are adaptable to high-throughput and/or automated operation.

Nucleic-acid-based diagnostic methods: The invention provides methods for altered levels or sequences of Gene 216 nucleic acids in a
10 sample, such as in a biological sample, which methods comprise the steps of: 1) contacting a sample suspected to contain a disease-associated nucleic acid with one or more disease-associated nucleic acid probes under conditions in which hybrids can form between any of the probes and disease-associated nucleic acid in the sample; and 2) detecting any hybrids
15 formed in step (1) using any suitable means known in the art, wherein the detection of hybrids indicates the presence of the disease-associated nucleic acid in the sample. To detect disease-associated nucleic acids present in low levels in biological samples, it may be necessary to amplify the disease-associated sequences or the hybridization signal as part of the
20 diagnostic assay. Techniques for amplification are known to those of skill in the art.

 The presence of Gene 216 polynucleotide sequences can be detected by DNA-DNA or DNA-RNA hybridization, or by amplification using probes or primers comprising at least a portion of a Gene 216
25 polynucleotide, or a sequence complementary thereto. In particular, nucleic acid amplification-based assays can use Gene 216 oligonucleotides or oligomers to detect transformants containing Gene 216 DNA or RNA. Gene 216 nucleic acids useful as probes in diagnostic methods include oligonucleotides at least 15 nucleotides in length, preferably at least 20
30 nucleotides in length, and most preferably at least 25-55 nucleotides in length, that hybridize specifically with Gene 216 nucleic acids.

Several methods can be used to produce specific probes for Gene 216 polynucleotides. For example, labeled probes can be produced by oligo-labeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, Gene 216 polynucleotide sequences (e.g., SEQ ID
5 NO:1 or SEQ ID NO:6), or any portions or fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase, such as T7, T3, or SP(6) and labeled nucleotides. These procedures may be conducted using
10 a variety of commercially available kits (e.g., from Amersham-Pharmacia; Promega Corp.; and U.S. Biochemical Corp., Cleveland, OH). Suitable reporter molecules or labels which may be used include radionucleotides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

15 A sample to be analyzed, such as, for example, a tissue sample (e.g., hair or buccal cavity) or body fluid sample (e.g., blood or saliva), may be contacted directly with the nucleic acid probes. Alternatively, the sample may be treated to extract the nucleic acids contained therein. It will be understood that the particular method used to extract DNA will depend on
20 the nature of the biological sample. The resulting nucleic acid from the sample may be subjected to gel electrophoresis or other size separation techniques, or, the nucleic acid sample may be immobilized on an appropriate solid matrix without size separation.

Kits suitable for nucleic acid-based diagnostic applications typically
25 include the following components:

- (1) Probe DNA: The probe DNA may be prelabeled; alternatively, the probe DNA may be unlabeled and the ingredients for labeling may be included in the kit in separate containers; and
- (2) Hybridization reagents: The kit may also contain other suitably
30 packaged reagents and materials needed for the particular hybridization protocol, including solid-phase matrices, if applicable, and standards.

In cases where a disease condition is suspected to involve an alteration of the Gene 216 nucleotide sequence, specific oligonucleotides may be constructed and used to assess the level of disease mRNA in cells affected or other tissue affected by the disease. For example, PCR can be
5 used to test whether a person has a disease-related polymorphism (i.e., mutation).

For PCR analysis, Gene 216 oligonucleotides may be chemically synthesized, generated enzymatically, or produced from a recombinant source. Oligomers will preferably comprise two nucleotide sequences, one
10 with a sense orientation (5' → 3') and another with an antisense orientation (3' → 5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA
15 sequences.

In accordance with PCR analysis, two oligonucleotides are synthesized by standard methods or are obtained from a commercial supplier of custom-made oligonucleotides. The length and base composition are determined by standard criteria using the Oligo 4.0 primer
20 Picking program (W. Rychlik, 1992; available from Molecular Biology Insights, Inc., Cascade, CO). One of the oligonucleotides is designed so that it will hybridize only to the disease gene DNA under the PCR conditions used. The other oligonucleotide is designed to hybridize a segment of genomic DNA such that amplification of DNA using these oligonucleotide
25 primers produces a conveniently identified DNA fragment. Samples may be obtained from hair follicles, whole blood, or the buccal cavity. The DNA fragment generated by this procedure is sequenced by standard techniques.

In one particular aspect, Gene 216 oligonucleotides can be used to perform Genetic Bit Analysis (GBA) of Gene 216 in accordance with
30 published methods (T.T. Nikiforov et al., 1994, *Nucleic Acids Res.* **22**(20):4167-75; T.T. Nikiforov et al., 1994, *PCR Methods Appl.* **3**(5):285-

91). In PCR-based GBA, specific fragments of genomic DNA containing the polymorphic site(s) are first amplified by PCR using one unmodified and one phosphorothioate-modified primer. The double-stranded PCR product is rendered single-stranded and then hybridized to immobilized oligonucleotide primer in wells of a multi-well plate. The primer is designed to anneal immediately adjacent to the polymorphic site of interest. The 3' end of the primer is extended using a mixture of individually labeled dideoxynucleoside triphosphates. The label on the extended base is then determined. Preferably, GBA is performed using semi-automated ELISA or biochip formats (see, e.g., S.R. Head et al., 1997, *Nucleic Acids Res.* **25**(24):5065-71; T.T. Nikiforov et al., 1994, *Nucleic Acids Res.* **22**(20):4167-75).

Other amplification techniques besides PCR may be used as alternatives, such as ligation-mediated PCR or techniques involving Q-beta replicase (Cahill et al., 1991, *Clin. Chem.*, **37**(9):1482-5). Products of amplification can be detected by agarose gel electrophoresis, quantitative hybridization, or equivalent techniques for nucleic acid detection known to one skilled in the art of molecular biology (Sambrook et al., 1989). Other alterations in the disease gene may be diagnosed by the same type of amplification-detection procedures, by using oligonucleotides designed to contain and specifically identify those alterations.

Gene 216 polynucleotides may also be used to detect and quantify levels of Gene 216 mRNA in biological samples in which altered expression of Gene 216 polynucleotide may be correlated with disease. These diagnostic assays may be used to distinguish between the absence, presence, increase, and decrease of Gene 216 mRNA levels, and to monitor regulation of Gene 216 polynucleotide levels during therapeutic treatment or intervention. For example, Gene 216 polynucleotide sequences, or fragments, or complementary sequences thereof, can be used in Southern or Northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or biochip assays utilizing fluids or tissues from patient biopsies to detect the status of, e.g., levels or

overexpression of Gene 216, or to detect altered Gene 216 expression. Such qualitative or quantitative methods are well known in the art (G.H. Keller and M.M. Manak, 1993, *DNA Probes*, 2nd Ed, Macmillan Publishers Ltd., England; D.W. Dieffenbach and G. S. Dveksler, 1995, *PCR Primer: A Laboratory*
5 *Manual*, Cold Spring Harbor Press, Plainview, NY; B.D. Hames and S.J. Higgins, 1985, *Gene Probes 1, 2*, IRL Press at Oxford University Press, Oxford, England).

Methods suitable for quantifying the expression of Gene 216 include radiolabeling or biotinylating nucleotides, co-amplification of a control nucleic
10 acid, and standard curves onto which the experimental results are interpolated (P.C. Melby et al., 1993, *J. Immunol. Methods* **159**:235-244; and C. Duplaa et al., 1993, *Anal. Biochem.* 229-236). The speed of quantifying multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric
15 or colorimetric response gives rapid quantification.

In accordance with these methods, the specificity of the probe, i.e., whether it is made from a highly specific region (e.g., at least 8 to 10 or 12 or 15 contiguous nucleotides in the 5' regulatory region), or a less specific region (e.g., especially in the 3' coding region), and the stringency of the hybridization
20 or amplification (e.g., high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding the Gene 216 polypeptide, alleles thereof, or related sequences.

In a particular aspect, a Gene 216 nucleic acid sequence, or a sequence complementary thereto, or fragment thereof, may be useful in
25 assays that detect Gene 216-related diseases such as asthma. The Gene 216 polynucleotide can be labeled by standard methods, and added to a biological sample from a subject under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample can be washed and the signal is quantified and compared with a standard value. If the
30 amount of signal in the test sample is significantly altered from that of a comparable negative control (normal) sample, the altered levels of Gene 216

nucleotide sequence can be correlated with the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular prophylactic or therapeutic regimen in animal studies, in clinical trials, or for an individual patient.

- 5 To provide a basis for the diagnosis of a disease associated with altered expression of Gene 216, a normal or standard profile for expression is established. This may be accomplished by incubating biological samples taken from normal subjects, either animal or human, with a sequence complementary to the Gene 216 polynucleotide, or a fragment thereof, under conditions
10 suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are
15 symptomatic for the disease. Deviation between standard and subject (patient) values is used to establish the presence of the condition.

Once the disease is diagnosed and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that which is
20 observed in a normal individual. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to diseases such as asthma, the presence of an abnormal amount of Gene 216 transcript in a biological sample (e.g., body fluid, cells,
25 tissues, or cell or tissue extracts) from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier, thereby preventing the
30 development or further progression of the disease.

Microarrays: In another embodiment of the present invention, oligonucleotides, or longer fragments derived from the Gene 216 polynucleotide sequence described herein may be used as targets in a microarray (e.g., biochip) system. The microarray can be used to monitor the expression level of large numbers of genes simultaneously (to produce a transcript image), and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disease, to diagnose disease, and to develop and monitor the activities of therapeutic or prophylactic agents. Preparation and use of microarrays have been described in WO 95/11995 to Chee et al.; D.J. Lockhart et al., 1996, *Nature Biotechnology* **14**:1675-1680; M. Schena et al., 1996, *Proc. Natl. Acad. Sci. USA* **93**:10614-10619; U.S. Patent No. 6,015,702 to P. Lal et al; J. Worley et al., 2000, *Microarray Biochip Technology*, M. Schena, ed., Biotechniques Book, Natick, MA, pp. 65-86; Y.H. Rogers et al., 1999, *Anal. Biochem.* **266**(1):23-30; S.R. Head et al., 1999, *Mol. Cell. Probes.* **13**(2):81-7; S.J. Watson et al., 2000, *Biol. Psychiatry* **48**(12):1147-56.

In one application of the present invention, microarrays containing arrays of Gene 216 polynucleotide sequences can be used to measure the expression levels of Gene 216 in an individual. In particular, to diagnose an individual with a Gene 216-related condition or disease, a sample from a human or animal (containing nucleic acids, e.g., mRNA) can be used as a probe on a biochip containing an array of Gene 216 polynucleotides (e.g., DNA) in decreasing concentrations (e.g., 1 ng, 0.1 ng, 0.01 ng, etc.). The test sample can be compared to samples from diseased and normal samples. Biochips can also be used to identify Gene 216 mutations or polymorphisms in a population, including but not limited to, deletions, insertions, and mismatches. For example, mutations can be identified by: 1) placing Gene 216 polynucleotides of this invention onto a biochip; 2) taking a test sample (containing, e.g., mRNA) and adding the sample to the biochip; 3) determining if the test samples hybridize to the Gene 216 polynucleotides attached to the

chip under various hybridization conditions (see, e.g., V.R. Chechetkin et al., 2000, *J. Biomol. Struct. Dyn.* **18**(1):83-101). Alternatively microarray sequencing can be performed (see, e.g., E.P. Diamandis, 2000, *Clin. Chem.* **46**(10):1523-5).

5 Chromosome mapping: In another application of this invention, the Gene 216 nucleic acid sequence, or a complementary sequence, or fragment thereof, can be used as probes which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to human artificial
10 chromosome constructions (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial PI constructions, or single chromosome cDNA libraries (see C.M. Price, 1993, *Blood Rev.*, 7:127-134 and by B.J. Trask, 1991, *Trends Genet.* 7:149-154).

15 In another of its aspects, the invention relates to a diagnostic kit for detecting Gene 216 polynucleotide or polypeptide as it relates to a disease or susceptibility to a disease, particularly asthma. Also related is a diagnostic kit that can be used to detect or assess asthma conditions. Such kits comprise one or more of the following:

20 (a) a Gene 216 polynucleotide, preferably the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:6, or a fragment thereof; or

 (b) a nucleotide sequence complementary to that of (a); or

 (c) a Gene 216 polypeptide, preferably the polypeptide of SEQ ID NO:4, or a fragment thereof; or

25 (d) an antibody to a Gene 216 polypeptide, preferably to the polypeptide of SEQ ID NO:4, or an antibody bindable fragment thereof. It will be appreciated that in any such kits, (a), (b), (c), or (d) may comprise a substantial component and that instructions for use can be included. The kits may also contain peripheral reagents such as buffers, stabilizers, etc.

30 The present invention also includes a test kit for genetic screening that can be utilized to identify mutations in Gene 216. By identifying patients with mutated Gene 216 DNA and comparing the mutation to a database that

contains known mutations in Gene 216 and a particular condition or disease, identification and/or confirmation of, a particular condition or disease can be made. Accordingly, such a kit would comprise a PCR-based test that would involve transcribing the patients mRNA with a specific primer, and amplifying

5 the resulting cDNA using another set of primers. The amplified product would be detectable by gel electrophoresis and could be compared with known standards for Gene 216. Preferably, this kit would utilize a patient's blood, serum, or saliva sample, and the DNA would be extracted using standard techniques. Primers flanking a known mutation would then be used to amplify

10 a fragment of Gene 216. The amplified piece would then be sequenced to determine the presence of a mutation.

Genomic Screening: The use of polymorphic genetic markers linked to the Gene 216 gene is very useful in predicting susceptibility to the diseases genetically linked to 20p13-p12. Similarly, the identification of polymorphic

15 genetic markers within the Gene 216 gene will allow the identification of specific allelic variants that are in linkage disequilibrium with other genetic lesions that affect one of the disease states discussed herein including respiratory disorders, obesity, and inflammatory bowel disease. SSCP (see below) allows the identification of polymorphisms within the genomic and

20 coding region of the disclosed gene. The present invention provides sequences for primers that can be used identify exons that contain SNPs, as well as sequences for primers that can be used to identify the sequence change. This information can be used to identify additional SNPs in accordance with the methods disclosed herein. Suitable methods for genomic

25 screening have also been described by, e.g., Sheffield et al., 1995, *Genet.*, 4:1837-1844; LeBlanc-Straceski et al., 1994, *Genomics*, 19:341-9; Chen et al., 1995, *Genomics*, 25:1-8. In employing these methods, the disclosed reagents can be used to predict the risk for disease (e.g., respiratory disorders, obesity, and inflammatory bowel disease) in a population or individual.

30 **Therapeutics**

The present invention provides methods of screening for drugs

comprising contacting such an agent with a novel protein of this invention or fragment thereof and assaying 1) for the presence of a complex between the agent and the protein or fragment, or 2) for the presence of a complex between the protein or fragment and a ligand, by methods well known in the art. In such competitive binding assays the novel protein or fragment is typically labeled. Free protein or fragment is separated from that present in a protein:protein complex, and the amount of free (i.e., uncomplexed) label is a measure of the binding of the agent being tested to Gene 216 protein or its interference with protein ligand binding, respectively.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of specifically binding the Gene 216 protein compete with a test compound for binding to the Gene 216 protein or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants of a Gene 216 protein.

The goal of rational drug design is to produce structural analogs of biologically active proteins of interest or of small molecules with which they interact (e.g., agonists, antagonists, inhibitors) in order to fashion drugs which are, for example, more active or stable forms of the protein, or which, e.g., enhance or interfere with the function of a protein *in vivo* (see, e.g., Hodgson, 1991, *Bio/Technology*, **9**:19-21). In one approach, one first determines the three-dimensional structure of a protein of interest or, for example, of the Gene 216 receptor or ligand complex, by x-ray crystallography, by computer modeling or most typically, by a combination of approaches. Less often, useful information regarding the structure of a protein may be gained by modeling based on the structure of homologous proteins. An example of rational drug design is the development of HIV protease inhibitors (Erickson et al., 1990, *Science*, **249**:527-533). In addition, peptides (e.g., Gene 216 protein) are analyzed by an alanine scan (Wells, 1991, *Methods in Enzymol.*, **202**:390-411). In this technique, an amino acid residue is replaced by Ala, and its effect on the peptide's activity

is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

It is also possible to isolate a target-specific antibody, selected by a functional assay, and then to solve its crystal structure. In principle, this approach yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original Gene 216 protein. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacore.

Thus, one may design drugs which result in, for example, altered Gene 216 protein activity or stability or which act as inhibitors, agonists, antagonists, etc. of Gene 216 protein activity. By virtue of the availability of cloned Gene 216 gene sequences, sufficient amounts of the Gene 216 protein may be made available to perform such analytical studies as x-ray crystallography. In addition, the knowledge of the Gene 216 polypeptide sequence will guide those employing computer-modeling techniques in place of, or in addition to x-ray crystallography.

In another aspect of the present invention, cells and animals that carry the Gene 216 gene or an analog thereof can be used as model systems to study and test for substances that have potential as therapeutic agents. After a test substance is administered to animals or applied to the cells, the phenotype of the animals/cells can be determined.

In yet another aspect of this invention, antibodies that specifically react with Gene 216 polypeptide or peptides derived therefrom can be used as therapeutics. In particular, anti-Gene 216 antibodies can be used to block the Gene 216 activity. Anti-Gene 216 antibodies or fragments thereof can be formulated as pharmaceutical compositions and administered to a subject. It is noted that antibody-based therapeutics produced from non-human sources

can cause an undesired immune response in human subjects. To minimize this problem, chimeric antibody derivatives can be produced. Chimeric antibodies combine a non-human animal variable region with a human constant region. Chimeric antibodies can be constructed according to methods known in the art (see Morrison et al., 1985, *Proc. Natl. Acad. Sci. USA*, 81:6851; Takeda et al., 1985, *Nature* 314:452; U.S. Patent No. 4,816,567 of Cabilly et al.; U.S. Patent No. 4,816,397 of Boss et al.; European Patent Publication EP 171496; EP 0173494; United Kingdom Patent GB 2177096B). In addition, antibodies can be further "humanized" by any of the techniques known in the art, (e.g., Teng et al., 1983, *Proc. Natl. Acad. Sci. USA* 80:7308-7312; Kozbor et al., 1983, *Immunology Today* 4: 7279; Olsson et al., 1982, *Meth. Enzymol.* 92:3-16; International Patent Application WO92/06193; EP 0239400). Humanized antibodies can also be obtained from commercial sources (e.g., Scotgen Limited, Middlesex, Great Britain). Immunotherapy with a humanized antibody may result in increased long-term effectiveness for the treatment of chronic disease situations or situations requiring repeated antibody treatments.

In one embodiment, compositions (e.g., pharmaceutical compositions) for use with the present invention comprise metalloprotease inhibitors, or analogs or derivatives thereof. Non-limiting examples of metalloprotease inhibitors include: 1) naturally occurring inhibitors, e.g., oprin (J.J. Catanese and L.F. Kress, 1992, *Biochemistry* 31:410-418; HSF (Y. Yamakawa and T. Omori-Satoh, 1992, *J. Biochem.* 112:583-589); erinacin (D. Mebs et al., 1996, *Toxicon* 34:1313-1316; Omori-Satoh et al., 2000, *Toxicon* 38:1561-1580); DM40 and DM43 (A.G. Neves-Ferreira et al., 2000, *Biochem. Biophys. Acta.* 1473:309-320); citrate (B. Francis et al., 1992, *Toxicon* 30:1239-1246); TIMP-1 and TIMP-2 (R.V. Ward et al., 1991, *Biochem J.* 278, Pt 1:179-873); pyrophosphate (G.S. Makowski and M.L. Ramsby, 1999, *Inflammation* 23:333-360); proglutamyl peptides such as pyroGlu-Asn-Trp-OH and pyroGlu-Glu-Trp-OH (A. Robeva et al., 1991, *Biomed. Biochem. Acta.* 50:769-773); 2) peptide analogs and derivatives, e.g., 2-distereomeric furan-2-carbonylamino-3-

oxohexahydroindolizino[8,7-b]indole carboxylates (S. D'Alessio et al., 2001, *Eur. J. Med. Chem.* **36**:43-53); phosphonate and carboxylate derivatives of pyroGlu-Asn-Trp-OH (D'Alessio et al., 2001); POL 647 and POL 656 (F.X. Gomis-Ruth et al., 1998, *Prot. Sci.* **7**:283-292); cysteine-switches (K. Nomura and N. Suzuki, 1993, *FEBS Lett.* **321**:84-88); 3) hydroxamate compounds, e.g., batimastat/BB-94 (see, e.g., G.F. Beattie et al., 1998, *Clin. Cancer Res.* **8**:1899-1902); prinomastat/AG3340 (see, e.g., R. Scatena, 2000, *Expert Opin. Investig. Drugs* **9**:2159-2165); and 4) other inhibitors, e.g., ortho-substituted macrocyclic lactams (G.M. Ksander, 1997, *J. Med. Chem.* **40**:495-505); diketopiperazine (DKP) (A.K. Szardenings et al., 1998, *J. Med. Chem.* **41**(13):2194-200; alendronate/PCP (Makowski and Ramsby, 1999); and CT1746 (Z. An et al., 1997, *Clin. Exp. Metastasis* **15**:184-195).

In particular, the determined structures of metalloproteases and metalloprotease inhibitors can be used to devise Gene 216-targeted inhibitors (i.e., by rational drug design; see Szardenings et al., 1998). Structural information can be found in, e.g., C. Oefner et al., 2000, *J. Mol. Biol.* **296**(2):341-9; B. Wu et al., 2000, *J. Mol. Biol.* **295**(2):257-68; L. Chen et al., 1999, *J. Mol. Biol.* **293**(3):545-57; C. Fernandez-Catalan et al., 1998, *EMBO J.* **17**(17):5238-48; S. Arumugam et al., 1998, *Biochemistry* **37**(27):9650-7; Gohlke et al., 1996, *FEBS Lett.* **378**:126-130; Gomis-Ruth et al., 1998; F.X. Gomis-Ruth et al., 1993, *EMBO J.* **12**:4151-4157; F.X. Gomis-Ruth et al., 1996, *J. Mol. Biol.* **264**:556-566; K. Maskos et al., 1998, *Proc. Natl. Acad. Sci. USA* **95**(7):3408-12; F.X. Gomis-Ruth et al., 1997, *Nature* **389**:77-80; M. Betz et al., 1997, *Eur. J. Biochem.* **247**(1):356-63; B. Lovejoy et al., 1994, *Biochemistry* **33**(27):8207-17. Structures of zinc metalloproteases are also found in Molecular Modeling DataBase (MMDB) at the NCBI web site <http://www.ncbi.nlm.nih.gov:80/Structure/MMDB/mmdb.shtml> (e.g. Accession Nos. 1D5J, 1D8F, 1D7X, 1BSK, 2TLX, 1TLX, 1BUD, 1BSW, 1UEA, 4AIG, 3AIG, 2AIG, 1KUH, 1DTH, 1UMS, 1UMT, 7TLN, 6TMN, 5TMN, 5TLN, 4TMN, 4TLN, 3TMN, 2TMN, 1TMN, 1TLP, 1IAG, 1HYT, 1AST, 8TLN, 1THL). In an alternative approach, the binding specificity of TIMP proteins can be

engineered to produce inhibitors that specifically inactivate Gene 216 polypeptide (see, e.g., H. Nagase et al., 1999, *Ann. NY Acad. Sci.* **878**:1-11; G.S. Butler et al., 1999, *J. Biol. Chem.* **274**(29):20391-20396).

In another embodiment of the present invention, compositions (e.g., pharmaceutical compositions) for use with the present invention comprise disintegrin agonists, or analogs or derivatives thereof. The determined structures of disintegrin proteins and domains can be used to devise Gene 216 disintegrin-targeted agonists (i.e., by rational drug design). Such structural information can be found in R.A. Atkinson et al., 1994, *Int. J. Pept. Protein Res.* **43**:563-72; V. Saudek et al., 1991, *Eur. J. Biochem.* **202**:329-38; H. Minoux et al., 2000, *J. Comput. Aided Mol. Des.* **14**:317-27.

The present invention contemplates compositions comprising a Gene 216 polynucleotide, polypeptide, antibody, ligand (e.g., agonist, antagonist, or inhibitor), or fragments, variants, or analogs thereof, and a physiologically acceptable carrier, excipient, or diluent as described in detail herein. The present invention further contemplates pharmaceutical compositions useful in practicing the therapeutic methods of this invention. Preferably, a pharmaceutical composition includes, in admixture, a pharmaceutically acceptable excipient (carrier) and one or more of a Gene 216 polypeptide, polynucleotide, ligand, antibody, or fragment or variant thereof, as described herein, as an active ingredient. The preparation of pharmaceutical compositions that contain Gene 216-related reagents as active ingredients is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions, however, solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified. The active therapeutic ingredient is often mixed with excipients that are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the composition can contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH-buffering agents, which

enhance the effectiveness of the active ingredient.

A Gene 216 polypeptide, polynucleotide, ligand, antibody, or variant or fragment thereof can be formulated into the pharmaceutical composition as neutralized physiologically acceptable salt forms. Suitable salts include the acid addition salts (i.e., formed with the free amino groups of the polypeptide or antibody molecule) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed from the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The pharmaceutical compositions can be administered systemically by oral or parenteral routes. Non-limiting parenteral routes of administration include subcutaneous, intramuscular, intraperitoneal, intravenous, transdermal, inhalation, intranasal, intra-arterial, intrathecal, enteral, sublingual, or rectal. Intravenous administration, for example, can be performed by injection of a unit dose. The term "unit dose" when used in reference to a pharmaceutical composition of the present invention refers to physically discrete units suitable as unitary dosage for humans, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent; i.e., carrier, or vehicle.

In one particular embodiment of the present invention, the disclosed pharmaceutical compositions are administered via mucoactive aerosol therapy (see, e.g., M. Fuloria and B.K. Rubin, 2000, *Respir. Care* 45:868-873; I. Gonda, 2000, *J. Pharm. Sci.* 89:940-945; R. Dhand, 2000, *Curr. Opin. Pulm. Med.* 6(1):59-70; B.K. Rubin, 2000, *Respir. Care* 45(6):684-94; S. Suarez and A.J. Hickey, 2000, *Respir. Care* 45(6):652-66).

Pharmaceutical compositions are administered in a manner compatible with the dosage formulation, and in a therapeutically effective amount. The quantity to be administered depends on the subject to be treated, capacity of

the subject's immune system to utilize the active ingredient, and degree of modulation of Gene 216 activity desired. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner and are specific for each individual. However, suitable dosages may range from about
5 0.1 to 20, preferably about 0.5 to about 10, and more preferably one to several, milligrams of active ingredient per kilogram body weight of individual per day and depend on the route of administration. Suitable regimes for initial administration and booster shots are also variable, but are typified by an initial administration followed by repeated doses at one or more hour intervals by a
10 subsequent injection or other administration. Alternatively, continuous intravenous infusions sufficient to maintain concentrations of 10 nM to 10 μ M in the blood are contemplated. An exemplary pharmaceutical formulation comprises: Gene 216 antagonist or inhibitor (5.0 mg/ml); sodium bisulfite USP (3.2 mg/ml); disodium edetate USP (0.1 mg/ml); and water for injection q.s.a.d.
15 (1.0 ml). As used herein, "pg" means picogram, "ng" means nanogram, " μ g" means microgram, "mg" means milligram, " μ l" means microliter, "ml" means milliliter, and "l" means L.

For further guidance in preparing pharmaceutical formulations, see, e.g., Gilman et al. (eds), 1990, *Goodman and Gilman's: The Pharmacological*
20 *Basis of Therapeutics*, 8th ed., Pergamon Press; and *Remington's Pharmaceutical Sciences*, 17th ed., 1990, Mack Publishing Co., Easton, PA; Avis et al. (eds), 1993, *Pharmaceutical Dosage Forms: Parenteral Medications*, Dekker, New York; Lieberman et al. (eds), 1990, *Pharmaceutical Dosage Forms: Disperse Systems*, Dekker, New York.

25 Pharmacogenetics: The Gene 216 polypeptides and polynucleotides are also useful in pharmacogenetic analysis (i.e., the study of the relationship between an individual's genotype and that individual's response to a therapeutic composition or drug). See, e.g., M. Eichelbaum, 1996, *Clin. Exp. Pharmacol. Physiol.* **23**(10-11):983-985, and M.W. Linder, 1997, *Clin. Chem.*
30 **43**(2):254-266. The genotype of the individual can determine the way a therapeutic acts on the body or the way the body metabolizes the therapeutic.

Further, the activity of drug metabolizing enzymes affects both the intensity and duration of therapeutic activity. Differences in the activity or metabolism of therapeutics can lead to severe toxicity or therapeutic failure. Accordingly, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenetic studies in determining whether to administer a Gene 216 polypeptide, polynucleotide, analog, antagonist, inhibitor, or modulator, as well as tailoring the dosage and/or therapeutic or prophylactic treatment regimen.

In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions can be due to a single factor that alters the way the drug act on the body (altered drug action), or a factor that alters the way the body metabolizes the drug (altered drug metabolism). These conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy which results in haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

The discovery of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. The gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response. This has been demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme, ultra-rapid metabolizers fail to respond to standard doses. Recent

studies have determined that ultra-rapid metabolism is attributable to CYP2D6 gene amplification.

By analogy, genetic polymorphism or mutation may lead to allelic variants of Gene 216 in the population which have different levels of activity.

- 5 The Gene 216 polypeptides or polynucleotides thereby allow a clinician to ascertain a genetic predisposition that can affect treatment modality. In addition, genetic mutation or variants at other genes may potentiate or diminish the activity of Gene 216-targeted drugs. Thus, in a Gene 216-based treatment, polymorphism or mutation may give rise to individuals that are more or less
- 10 responsive to treatment. Accordingly, dosage would necessarily be modified to maximize the therapeutic effect within a given population containing the polymorphism. As an alternative to genotyping, specific polymorphic polypeptides or polynucleotides can be identified.

- To identify genes that modify Gene 216-targeted drug response, several
- 15 pharmacogenetic methods can be used. One pharmacogenomics approach, "genome-wide association", relies primarily on a high-resolution map of the human genome. This high-resolution map shows previously identified gene-related markers (e.g., a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which
- 20 has two variants). A high-resolution genetic map can then be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, a high-resolution map can be generated from a combination of some ten million
- 25 known single nucleotide polymorphisms (SNPs) in the human genome. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In this way, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be
- 30 common among such genetically similar individuals (see, e.g., D.R. Pfost et al., 2000, *Trends Biotechnol.* **18**(8):334-8).

As another example, the "candidate gene approach", can be used. According to this method, if a gene that encodes a drug target is known, all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As yet another example, a "gene expression profiling approach", can be used. This method involves testing the gene expression of an animal treated with a drug (e.g., a Gene 216 polypeptide, polynucleotide, analog, or modulator) to determine whether gene pathways related to toxicity have been turned on.

Information obtained from one of the approaches described herein can be used to establish a pharmacogenetic profile, which can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. A pharmacogenetic profile, when applied to dosing or drug selection, can be used to avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a Gene 216 polypeptide, polynucleotide, analog, antagonist, inhibitor, or modulator.

Gene 216 polypeptides or polynucleotides are also useful for monitoring therapeutic effects during clinical trials and other treatment. Thus, the therapeutic effectiveness of an agent that is designed to increase or decrease gene expression, polypeptide levels, or activity can be monitored over the course of treatment using the Gene 216 compositions or modulators. For example, monitoring can be performed by: 1) obtaining a pre-administration sample from a subject prior to administration of the agent; 2) detecting the level of expression or activity of the protein in the pre-administration sample; 3) obtaining one or more post-administration samples from the subject; 4) detecting the level of expression or activity of the polypeptide in the post-administration samples; 5) comparing the level of expression or activity of the polypeptide in the pre-administration sample with the polypeptide in the post-administration sample or samples; and 6) increasing or decreasing the

administration of the agent to the subject accordingly.

Gene Therapy: In recent years, significant technological advances have been made in the area of gene therapy for both genetic and acquired diseases (Kay et al., 1997, *Proc. Natl. Acad. Sci. USA*, **94**:12744-12746). Gene therapy
5 can be defined as the transfer of DNA for therapeutic purposes. Improvement in gene transfer methods has allowed for development of gene therapy protocols for the treatment of diverse types of diseases. Gene therapy has also taken advantage of recent advances in the identification of new therapeutic genes, improvement in both viral and non-viral gene delivery
10 systems, better understanding of gene regulation, and improvement in cell isolation and transplantation. Gene therapy would be carried out according to generally accepted methods as described by, for example, Friedman, 1991, *Therapy for Genetic Diseases*, Friedman, Ed., Oxford University Press, pages 105-121.

15 Vectors for introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector may be used. Methods for introducing DNA into cells such as electroporation, calcium phosphate co-precipitation, and viral transduction are known in the art, and the choice of method is within the competence of
20 one skilled in the art (Robbins (ed), 1997, *Gene Therapy Protocols*, Human Press, NJ). Cells transformed with a Gene 216 gene can be used as model systems to study chromosome 20 disorders and to identify drug treatments for the treatment of such disorders.

Gene transfer systems known in the art may be useful in the practice
25 of the gene therapy methods of the present invention. These include viral and non-viral transfer methods. A number of viruses have been used as gene transfer vectors, including polyoma, *i.e.*, SV40 (Madzak et al., 1992, *J. Gen. Virol.*, **73**:1533-1536), adenovirus (Berkner, 1992, *Curr. Top. Microbiol. Immunol.*, **158**:39-6; Berkner et al., 1988, *Bio Techniques*, **6**:616-
30 629; Gorziglia et al., 1992, *J. Virol.*, **66**:4407-4412; Quantin et al., 1992, *Proc. Natl. Acad. Sci. USA*, **89**:2581-2584; Rosenfeld et al., 1992, *Cell*,

68:143-155; Wilkinson et al., 1992, *Nucl. Acids Res.*, **20**:2233-2239; Stratford-Perricaudet et al., 1990, *Hum. Gene Ther.*, **1**:241-256), vaccinia virus (Mackett et al., 1992, *Biotechnology*, **24**:495-499), adeno-associated virus (Muzyczka, 1992, *Curr. Top. Microbiol. Immunol.*, **158**:91-123; Ohi et al., 1990, *Gene*, **89**:279-282), herpes viruses including HSV and EBV
5 (Margolskee, 1992, *Curr. Top. Microbiol. Immunol.*, **158**:67-90; Johnson et al., 1992, *J. Virol.*, **66**:2952-2965; Fink et al., 1992, *Hum. Gene Ther.*, **3**:11-19; Breakfield et al., 1987, *Mol. Neurobiol.*, **1**:337-371; Fresse et al., 1990, *Biochem. Pharmacol.*, **40**:2189-2199), and retroviruses of avian
10 (Brandyopadhyay et al., 1984, *Mol. Cell Biol.*, **4**:749-754; Petropoulos et al., 1992, *J. Virol.*, **66**:3391-3397), murine (Miller, 1992, *Curr. Top. Microbiol. Immunol.*, **158**:1-24; Miller et al., 1985, *Mol. Cell Biol.*, **5**:431-437; Sorge et al., 1984, *Mol. Cell Biol.*, **4**:1730-1737; Mann et al., 1985, *J. Virol.*, **54**:401-407), and human origin (Page et al., 1990, *J. Virol.*, **64**:5370-5276;
15 Buchschalcher et al., 1992, *J. Virol.*, **66**:2731-2739). Most human gene therapy protocols have been based on disabled murine retroviruses.

Non-viral gene transfer methods known in the art include chemical techniques such as calcium phosphate coprecipitation (Graham et al., 1973, *Virology*, **52**:456-467; Pellicer et al., 1980, *Science*, **209**:1414-1422),
20 mechanical techniques, for example microinjection (Anderson et al., 1980, *Proc. Natl. Acad. Sci. USA*, **77**:5399-5403; Gordon et al., 1980, *Proc. Natl. Acad. Sci. USA*, **77**:7380-7384; Brinster et al., 1981, *Cell*, **27**:223-231; Constantini et al., 1981, *Nature*, **294**:92-94), membrane fusion-mediated transfer via liposomes (Felgner et al., 1987, *Proc. Natl. Acad. Sci. USA*,
25 **84**:7413-7417; Wang et al., 1989, *Biochemistry*, **28**:9508-9514; Kaneda et al., 1989, *J. Biol. Chem.*, **264**:12126-12129; Stewart et al., 1992, *Hum. Gene Ther.*, **3**:267-275; Nabel et al., 1990, *Science*, **249**:1285-1288; Liri et al., 1992, *Circulation*, **83**:2007-2011), and direct DNA uptake and receptor-mediated DNA transfer (Wolff et al., 1990, *Science*, **247**:1465-1468; Wu et al., 1991, *BioTechniques*, **11**:474-485; Zenke et al., 1990, *Proc. Natl. Acad.*
30

- Sci. USA*, **87**:3655-3659; Wu et al., 1989, *J. Biol. Chem.*, **264**:16985-16987; Wolff et al., 1991, *BioTechniques*, **11**:474-485; Wagner et al., 1991, *Proc. Natl. Acad. Sci. USA*, **88**:4255-4259; Cotten et al., 1990, *Proc. Natl. Acad. Sci. USA*, **87**:4033-4037; Curiel et al., 1991, *Proc. Natl. Acad. Sci. USA*, **88**:8850-8854; Curiel et al., 1991, *Hum. Gene Ther.*, **3**:147-154).

In one approach, plasmid DNA is complexed with a polylysine-conjugated antibody specific to the adenovirus hexon protein, and the resulting complex is bound to an adenovirus vector. The trimolecular complex is then used to infect cells. The adenovirus vector permits efficient binding, internalization, and degradation of the endosome before the coupled DNA is damaged.

In another approach, liposome/DNA is used to mediate direct *in vivo* gene transfer. While in standard liposome preparations the gene transfer process is non-specific, localized *in vivo* uptake and expression have been reported in tumor deposits, for example, following direct *in situ* administration (Nabel, 1992, *Hum. Gene Ther.*, **3**:399-410).

Suitable gene transfer vectors possess a promoter sequence, preferably a promoter that is cell-specific and placed upstream of the sequence to be expressed. The vectors may also contain, optionally, one or more expressible marker genes for expression as an indication of successful transfection and expression of the nucleic acid sequences contained in the vector. In addition, vectors can be optimized to minimize undesired immunogenicity and maximize long-term expression of the desired gene product(s) (see Nabe, 1999, *Proc. Natl. Acad. Sci. USA* **96**:324-326). Moreover, vectors can be chosen based on cell-type that is targeted for treatment. Notably, gene transfer therapies have been initiated for the treatment of various pulmonary diseases (see, e.g., M.J. Welsh, 1999, *J. Clin. Invest.* **104**(9):1165-6; D.L. Ennist, 1999, *Trends Pharmacol. Sci.* **20**:260-266; S.M. Albelda et al., 2000, *Ann. Intern. Med.* **132**:649-660; E. Alton and C. Kitson C., 2000, *Expert Opin. Investig. Drugs.* **9**(7):1523-35).

Illustrative examples of vehicles or vector constructs for transfection or

infection of the host cells include replication-defective viral vectors, DNA virus or RNA virus (retrovirus) vectors, such as adenovirus, herpes simplex virus and adeno-associated viral vectors. Adeno-associated virus vectors are single stranded and allow the efficient delivery of multiple copies of nucleic acid to the cell's nucleus. Preferred are adenovirus vectors. The vectors will normally be substantially free of any prokaryotic DNA and may comprise a number of different functional nucleic acid sequences. An example of such functional sequences may be a DNA region comprising transcriptional and translational initiation and termination regulatory sequences, including promoters (e.g., strong promoters, inducible promoters, and the like) and enhancers which are active in the host cells. Also included as part of the functional sequences is an open reading frame (polynucleotide sequence) encoding a protein of interest. Flanking sequences may also be included for site-directed integration. In some situations, the 5'-flanking sequence will allow homologous recombination, thus changing the nature of the transcriptional initiation region, so as to provide for inducible or non-inducible transcription to increase or decrease the level of transcription, as an example.

In general, the encoded and expressed Gene 216 polypeptide may be intracellular, i.e., retained in the cytoplasm, nucleus, or in an organelle, or may be secreted by the cell. For secretion, the natural signal sequence present in Gene 216 may be retained. When the polypeptide or peptide is a fragment of a Gene 216 protein, a signal sequence may be provided so that, upon secretion and processing at the processing site, the desired protein will have the natural sequence. Specific examples of coding sequences of interest for use in accordance with the present invention include the Gene polypeptide coding sequences, e.g., SEQ ID NO:4.

As previously mentioned, a marker may be present for selection of cells containing the vector construct. The marker may be an inducible or non-inducible gene and will generally allow for positive selection under induction, or without induction, respectively. Examples of marker genes include neomycin, dihydrofolate reductase, glutamine synthetase, and the like. The

vector employed will generally also include an origin of replication and other genes that are necessary for replication in the host cells, as routinely employed by those having skill in the art. As an example, the replication system comprising the origin of replication and any proteins associated with replication encoded by a particular virus may be included as part of the construct. The replication system must be selected so that the genes encoding products necessary for replication do not ultimately transform the cells. Such replication systems are represented by replication-defective adenovirus (see G. Acsadi et al., 1994, *Hum. Mol. Genet.* 3:579-584) and by Epstein-Barr virus. Examples of replication defective vectors, particularly, retroviral vectors that are replication defective, are BAG, (see Price et al., 1987, *Proc. Natl. Acad. Sci. USA*, 84:156; Sanes et al., 1986, *EMBO J.*, 5:3133). It will be understood that the final gene construct may contain one or more genes of interest, for example, a gene encoding a bioactive metabolic molecule. In addition, cDNA, synthetically produced DNA or chromosomal DNA may be employed utilizing methods and protocols known and practiced by those having skill in the art.

According to one approach for gene therapy, a vector encoding a Gene 216 polypeptide is directly injected into the recipient cells (*in vivo* gene therapy). Alternatively, cells from the intended recipients are explanted, genetically modified to encode a Gene 216 polypeptide, and reimplanted into the donor (*ex vivo* gene therapy). An *ex vivo* approach provides the advantage of efficient viral gene transfer, which is superior to *in vivo* gene transfer approaches. In accordance with *ex vivo* gene therapy, the host cells are first transfected with engineered vectors containing at least one gene encoding a Gene 216 polypeptide, suspended in a physiologically acceptable carrier or excipient such as saline or phosphate buffered saline, and the like, and then administered to the host. The desired gene product is expressed by the injected cells, which thus introduce the gene product into the host. The introduced gene products can thereby be utilized to treat or ameliorate a disorder that is related to altered levels of Gene 216 (e.g., asthma).

Animal Models

Gene 216 polynucleotides can be used to generate genetically altered non-human animals or human cell lines. Any non-human animal can be used; however typical animals are rodents, such as mice, rats, or guinea pigs.

- 5 Genetically engineered animals or cell lines can carry a gene that has been altered to contain deletions, substitutions, insertions, or modifications of the polynucleotide sequence (e.g., exon sequence). Such alterations may render the gene nonfunctional, (i.e., a null mutation) producing a "knockout" animal or cell line. In addition, genetically engineered animals can carry one or more
- 10 exogenous or non-naturally occurring genes, i.e., "transgenes", that are derived from different organisms (e.g., humans), or produced by synthetic or recombinant methods. Genetically altered animals or cell lines can be used to study Gene 216 function, regulation, and treatments for Gene 216-related diseases. In particular, knockout animals and cell lines can be used to
- 15 establish animal models and *in vitro* models for Gene 216-related illnesses, respectively. In addition, transgenic animals expressing human Gene 216 can be used in drug discovery efforts.

- A "transgenic animal" is any animal containing one or more cells bearing genetic information altered or received, directly or indirectly, by
- 20 deliberate genetic manipulation at a subcellular level, such as by targeted recombination or microinjection or infection with recombinant virus. The term "transgenic animal" is not intended to encompass classical cross-breeding or *in vitro* fertilization, but rather is meant to encompass animals in which one or more cells are altered by, or receive, a recombinant DNA molecule. This
- 25 recombinant DNA molecule may be specifically targeted to a defined genetic locus, may be randomly integrated within a chromosome, or it may be extrachromosomally replicating DNA.

- Transgenic animals can be selected after treatment of germline cells or zygotes. For example, expression of an exogenous Gene 216 gene or a
- 30 variant can be achieved by operably linking the gene to a promoter and optionally an enhancer, and then microinjecting the construct into a zygote

(see, e.g., Hogan et al., *Manipulating the Mouse Embryo, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). Such treatments include insertion of the exogenous gene and disrupted homologous genes. Alternatively, the gene(s) of the animals may be disrupted by insertion
5 or deletion mutation of other genetic alterations using conventional techniques (see, e.g., Capecchi, 1989, *Science*, **244**:1288; Valancuis et al., 1991, *Mol. Cell Biol.*, **11**:1402; Hasty et al., 1991, *Nature*, **350**:243; Shinkai et al., 1992, *Cell*, **68**:855; Mombaerts et al., 1992, *Cell*, **68**:869; Philpott et al., 1992, *Science*, **256**:1448; Snouwaert et al., 1992, *Science*, **257**:1083; Donehower et al., 1992,
10 *Nature*, **356**:215).

In one aspect of the invention, Gene 216 knockout mice can be produced in accordance with well-known methods (see, e.g., M.R. Capecchi, 1989, *Science*, **244**:1288-1292; P. Li et al., 1995, *Cell* **80**:401-411; L.A. Galli-Taliadoros et al., 1995, *J. Immunol. Methods* **181**(1):1-15; C.H. Westphal et al.,
15 1997, *Curr. Biol.* **7**(7):530-3; S.S. Cheah et al., 2000, *Methods Mol. Biol.* **136**:455-63). The disclosed murine Gene 216 genomic clone can be used to prepare a Gene 216 targeting construct that can disrupt Gene 216 in the mouse by homologous recombination at the Gene 216 chromosomal locus. The targeting construct can comprise a disrupted or deleted Gene 216
20 sequence that inserts in place of the functioning portion of the native mouse gene. For example, the construct can contain an insertion in the Gene 216 protein-coding region.

Preferably, the targeting construct contains markers for both positive and negative selection. The positive selection marker allows the selective
25 elimination of cells that lack the marker, while the negative selection marker allows the elimination of cells that carry the marker. In particular, the positive selectable marker can be an antibiotic resistance gene, such as the neomycin resistance gene, which can be placed within the coding sequence of Gene 216 to render it non-functional, while at the same time rendering the construct
30 selectable. The herpes simplex virus thymidine kinase (HSV tk) gene is an example of a negative selectable marker that can be used as a second marker

to eliminate cells that carry it. Cells with the HSV tk gene are selectively killed in the presence of gangcyclovir. As an example, a positive selection marker can be positioned on a targeting construct within the region of the construct that integrates at the Gene 216 locus. The negative selection marker can be
5 positioned on the targeting construct outside the region that integrates at the Gene 216 locus. Thus, if the entire construct is present in the cell, both positive and negative selection markers will be present. If the construct has integrated into the genome, the positive selection marker will be present, but the negative selection marker will be lost.

10 The targeting construct can be employed, for example, in embryonal stem cell (ES). ES cells may be obtained from pre-implantation embryos cultured *in vitro* (M.J. Evans et al., 1981, *Nature* **292**:154-156; M.O. Bradley et al., 1984, *Nature* **309**:255-258; Gossler et al., 1986, *Proc. Natl. Acad. Sci. USA* **83**:9065-9069; Robertson et al., 1986, *Nature* **322**:445-448; S. A. Wood et al.,
15 1993, *Proc. Natl. Acad. Sci. USA* **90**:4582-4584). Targeting constructs can be efficiently introduced into the ES cells by standard techniques such as DNA transfection or by retrovirus-mediated transduction. Following this, the transformed ES cells can be combined with blastocysts from a non-human animal. The introduced ES cells colonize the embryo and contribute to the
20 germ line of the resulting chimeric animal (R. Jaenisch, 1988, *Science* **240**:1468-1474). The use of gene-targeted ES cells in the generation of gene-targeted transgenic mice has been previously described (Thomas et al., 1987, *Cell* **51**:503-512) and is reviewed elsewhere (Frohman et al., 1989, *Cell* **56**:145-147; Capecchi, 1989, *Trends in Genet.* **5**:70-76; Baribault et al., 1989, *Mol. Biol. Med.* **6**:481-492; Wagner, 1990, *EMBO J.* **9**:3025-3032; Bradley et al.,
25 1992, *Bio/Technology* **10**: 534-539).

Several methods can be used to select homologously recombined murine ES cells. One method employs PCR to screen pools of transformant cells for homologous insertion, followed by screening individual clones (Kim et al., 1988, *Nucleic Acids Res.* **16**:8887-8903; Kim et al., 1991, *Gene* **103**:227-
30 233). Another method employs a marker gene is constructed which will only

be active if homologous insertion occurs, allowing these recombinants to be selected directly (Sedivy et al., 1989, *Proc. Natl. Acad. Sci. USA* **86**:227-231).

For example, the positive-negative selection (PNS) method can be used as described above (see, e.g., Mansour et al., 1988, *Nature* **336**:348-352; 5 Capecchi, 1989, *Science* **244**:1288-1292; Capecchi, 1989, *Trends in Genet.* **5**:70-76). In particular, the PNS method is useful for targeting genes that are expressed at low levels.

The absence of functional Gene 216 in the knockout mice can be confirmed, for example, by RNA analysis, protein expression analysis, and 10 functional studies. For RNA analysis, RNA samples are prepared from different organs of the knockout mice and the Gene 216 transcript is detected in Northern blots using oligonucleotide probes specific for the transcript. For protein expression detection, antibodies that are specific for the Gene 216 polypeptide are used, for example, in flow cytometric analysis, 15 immunohistochemical staining, and activity assays. Alternatively, functional assays are performed using preparations of different cell types collected from the knockout mice.

Several approaches can be used to produce transgenic mice. In one approach, a targeting vector is integrated into ES cell by homologous 20 recombination, an intrachromosomal recombination event is used to eliminate the selectable markers, and only the transgene is left behind (A.L. Joyner et al., 1989, *Nature* **338**(6211):153-6; P. Hasty et al., 1991, *Nature* **350**(6315):243-6; V. Valancius and O. Smithies, 1991, *Mol. Cell Biol.* **11**(3):1402-8; S. Fiering et al., 1993, *Proc. Natl. Acad. Sci. USA* **90**(18):8469-73). In an alternative 25 approach, two or more strains are created; one strain contains the gene knocked-out by homologous recombination, while one or more strains contain transgenes. The knockout strain is crossed with the transgenic strain to produce new line of animals in which the original wild-type allele has been replaced (although not at the same site) with a transgene. Notably, knockout 30 and transgenic animals can be produced by commercial facilities (e.g., The Lerner Research Institute, Cleveland, OH; B&K Universal, Inc., Fremont, CA;

DNX Transgenic Sciences, Cranbury, NJ; Incyte Genomics, Inc., St. Louis, MO).

Transgenic animals (e.g., mice) containing a nucleic acid molecule which encodes human Gene 216, may be used as *in vivo* models to study the overexpression of Gene 216. Such animals can also be used in drug evaluation and discovery efforts to find compounds effective to inhibit or modulate the activity of Gene 216, such as for example compounds for treating respiratory disorders, diseases, or conditions. One having ordinary skill in the art can use standard techniques to produce transgenic animals which produce human Gene 216 polypeptide, and use the animals in drug evaluation and discovery projects (see, e.g., U.S. Patent No. 4,873,191 to Wagner; U.S. Patent No. 4,736,866 to Leder).

In another embodiment of the present invention, the transgenic animal can comprise a recombinant expression vector in which the nucleotide sequence that encodes human Gene 216 is operably linked to a tissue specific promoter whereby the coding sequence is only expressed in that specific tissue. For example, the tissue specific promoter can be a mammary cell specific promoter and the recombinant protein so expressed is recovered from the animal's milk.

In yet another embodiment of the present invention, a Gene 216 "knockout" can be produced by administering to the animal antibodies (e.g., neutralizing antibodies) that specifically recognize an endogenous Gene 216 polypeptide. The antibodies can act to disrupt function of the endogenous Gene 216 polypeptide, and thereby produce a null phenotype. In one specific example, an orthologous mouse Gene 216 polypeptide (e.g., SEQ ID NO:366) or peptide can be used to generate antibodies. These antibodies can be given to a mouse to knockout the function of the mouse Gene 216 ortholog.

In addition, non-mammalian organisms may be used to study Gene 216 and Gene 216-related diseases. For example, model organisms such as *C. elegans*, *D. melanogaster*, and *S. cerevisiae* may be used. Gene 216 homologues can be identified in these model organisms, and mutated or

deleted to produce a Gene 216-deficient strain. Human Gene 216 can then be tested for the ability to "complement" the Gene 216-deficient strain. Gene 216-deficient strains can also be used for drug screening. The study of Gene 216 homologs can facilitate the understanding of human Gene 216 biological function, and assist in the identification of binding proteins (e.g., agonists and antagonists).

Gene Identification

To identify genes in the region on 20p13-p12, a set of bacterial artificial chromosome(BAC) clones containing this chromosomal region was identified in accordance with the methods described herein. The BAC clones served as a template for genomic DNA sequencing and served as reagents for identifying coding sequences by direct cDNA selection. Genomic sequencing and direct cDNA selection methods were used to characterize DNA from 20p13-p12.

When one or more genes have been genetically localized to a specific chromosomal region, the gene(s) can be characterized at the molecular level by a series of steps that include: 1) cloning the entire region of DNA in a set of overlapping clones (physical mapping); 2) characterizing the gene(s) encoded by these clones by a combination of direct cDNA selection, exon trapping and DNA sequencing (gene identification); and 3) identifying mutations (i.e., SNPs) in the gene(s) by comparative DNA sequencing of affected and unaffected members of the kindred and/or in unrelated affected individuals and unrelated unaffected controls (mutation analysis).

Physical mapping is accomplished by screening libraries of human DNA cloned in vectors that are propagated in a host such as *E. coli*, using hybridization or PCR assays from unique molecular landmarks in the chromosomal region of interest. In accordance with the present invention, a physical map of the disorder region was generated by screening a library of human DNA cloned in BACs with a set overgo markers that had been previously mapped to chromosome 20p13-p12 by the efforts of the Human Genome Project. Overgos are unique molecular landmarks in the human genome that can be assayed by hybridization. The location of thousands of

overgos on the twenty-two autosomes and two sex chromosomes has been determined through the efforts of the Human Genome Project. For a positional cloning effort, the physical map is tied to the genetic map because the markers used for genetic mapping can also be used as overgos for physical mapping.

- 5 By screening a BAC library with a combination of overgos derived from genetic markers, genes, and random DNA fragments, a physical map comprised of overlapping clones representing all of the DNA in a chromosomal region of interest can be assembled.

- 10 BACs are cloning vectors for large (80 kilobase to 200 kilobase) segments of human or other DNA that are propagated in *E. coli*. To construct a physical map using BACs, a library of BAC clones is screened so that individual clones harboring the DNA sequence corresponding to a given overgo or set of overgos are identified. Throughout most of the human genome, the overgo markers are spaced approximately 20 to 50 kilobases apart, so that an
15 individual BAC clone typically contains at least two overgo markers. In addition, the BAC libraries that were screened contain enough cloned DNA to cover the human genome twelve times over. An individual overgo typically identifies more than one BAC clone. By screening a twelve-fold coverage BAC library with a series of overgo markers spaced approximately 50 kilobases
20 apart, a physical map consisting of a series of overlapping contiguous BAC clones, i.e., BAC "contigs," can be assembled for any region of the human genome. This map is closely tied to the genetic map because many of the overgo markers used to prepare the physical map are also genetic markers.

- 25 When constructing a physical map, it often happens that there are gaps in the overgo map of the genome that result in the inability to identify BAC clones that are overlapping in a given location. Typically, the physical map is first constructed from a set of overgos identified through the publicly available literature and World Wide Web resources. The initial map consists of several separate BAC contigs that are separated by gaps of unknown molecular
30 distance. To identify BAC clones that fill these gaps, it is necessary to develop new overgo markers from the ends of the clones on either side of the gap.

This is done by sequencing the terminal 200 to 300 base pairs of the BACs flanking the gap, and developing a PCR or hybridization based assay. If the terminal sequences are demonstrated to be unique within the human genome, then the new overgo can be used to screen the BAC library to identify
5 additional BACs that contain the DNA from the gap in the physical map. To assemble a BAC contig that covers a region the size of the disorder region (6,000,000 or more base pairs), it is necessary to develop new overgo markers from the ends of a number of clones.

After building a BAC contig, this set of overlapping clones serves as a
10 template for identifying the genes encoded in the chromosomal region. Gene identification can be accomplished by many methods. Three methods are commonly used: 1) a set of BACs selected from the BAC contig to represent the entire chromosomal region are sequenced, and computational methods are used to identify all of the genes; 2) the BACs from the BAC contig are used as
15 a reagent to clone cDNAs corresponding to the genes encoded in the region by a method termed direct cDNA selection; or 3) the BACs from the BAC contig are used to identify coding sequences by selecting for specific DNA sequence motifs in a procedure called exon trapping. Gene 216 was identified by methods (1) and (2) in accordance with the techniques disclosed herein.

20 To sequence the entire BAC contig representing the disorder region, a set of BACs can be chosen for subcloning into plasmid vectors and subsequent DNA sequencing of these subclones. Since the DNA cloned in the BACs represents genomic DNA, this sequencing is referred to as genomic sequencing to distinguish it from cDNA sequencing. To initiate the genomic
25 sequencing for a chromosomal region of interest, several non-overlapping BAC clones are chosen. DNA for each BAC clone is prepared, and the clones are sheared into random small fragments that are subsequently cloned into standard plasmid vectors such as pUC18. The plasmid clones are then grown to propagate the smaller fragments, and these are the templates for
30 sequencing. To ensure adequate coverage and sequence quality for the BAC DNA sequence, sufficient plasmid clones are sequenced to yield three-fold

coverage of the BAC clone. For example, if the BAC is 100 kilobases long, then phagemids are sequenced to yield 300 kilobases of sequence. Since the BAC DNA is randomly sheared prior to cloning in the phagemid vector, the 300 kilobases of raw DNA sequence can be assembled by computational methods
5 into overlapping DNA sequences termed sequence contigs. For the purposes of initial gene identification by computational methods, three-fold coverage of each BAC is sufficient to yield twenty to forty sequence contigs of 1000 base pairs to 20,000 base pairs.

In accordance with the present invention, the "seed" BACs from the
10 BAC contig in the disorder region were sequenced. The sequence of the "seed" BACs was then used to identify minimally overlapping BACs from the contig, and these were subsequently sequenced. In this manner, the entire candidate region can be sequenced, with several small sequence gaps left in each BAC. This sequence serves as the template for computational gene
15 identification. In one approach, genes can be identified by comparing the sequence of BAC contig to publicly available databases of cDNA and genomic sequences, e.g. UniGene, dbEST, EMBL nucleotide database, GenBank, and the DNA Database of Japan (DDBJ). The BAC DNA sequence can also be translated into protein sequence, and the protein sequence can be used to
20 search publicly available protein databases, e.g., GenPept, EMBL protein database, Protein Information Resource (PIR), Protein Data Bank (PDB), and SWISS-PROT. These comparisons are typically done using the BLAST family of computer algorithms and programs (Altschul et al., 1990, *J. Mol. Biol.*, 215:403-410; Altschul et al, 1997, *Nucl. Acids Res.*, 25:3389-3402).

25 For nucleotide queries, BLASTN, BLASTX, and TBLASTX can be used. BLASTN compares a nucleotide query sequence with a nucleotide sequence database; BLASTX compares a nucleotide query sequence translated in all reading frames against a protein sequence database; TBLASTX compares the six-frame translations of a nucleotide query sequence against the six-frame
30 translations of a nucleotide sequence database. For protein queries, BLASTP and TBLASTN can be used. BLASTP compares a protein query sequence

with a protein sequence database; TBLASTN compares a protein query sequence against a nucleotide sequence database dynamically translated in all reading frames.

Additionally, computer algorithms such as MZEF (Zhang, 1997, *Proc. Natl. Acad. Sci. USA* **94**:565-568), GRAIL (Uberbacher et al., 1996, *Methods Enzymol.*, **266**:259-281), and Genscan (Burge and Karlin, 1997, *J. Mol. Biol.*, **268**:78-94) can be used to predict the location of exons in the sequence based on the presence of specific DNA sequence motifs that are common to all exons, as well as the presence of codon usage typical of human protein encoding sequences.

In addition to identifying genes by computational methods, genes can be identified by direct cDNA selection (Del Mastro and Lovett, 1996, *Methods in Molecular Biology*, Humana Press Inc., NJ). In direct cDNA selection, cDNA pools from tissues of interest are prepared, and BACs from the candidate region are used in a liquid hybridization assay to capture the cDNAs which base pair to coding regions in the BAC. In the methods described herein, the cDNA pools were created from several different tissues by random priming and oligo dT priming the first strand cDNA from poly A⁺ RNA, synthesizing the second-strand cDNA by standard methods, and adding linkers to the ends of the cDNA fragments. In this approach, the linkers are used to amplify the cDNA pools of BAC clones from the disorder region identified by screening a BAC library. The amplified products are then used as a template for initiating DNA synthesis to create a biotin labeled copy of BAC DNA. Following this, the biotin labeled copy of the BAC DNA is denatured and incubated with an excess of the PCR amplified, linkered cDNA pools which have also been denatured. The BAC DNA and cDNA are allowed to anneal in solution, and heteroduplexes between the BAC and the cDNA are isolated using streptavidin coated magnetic beads. The cDNAs that are captured by the BAC are then amplified using primers complimentary to the linker sequences, and the hybridization/selection process is repeated for a second round. After two rounds of direct cDNA selection, the cDNA fragments are cloned, and a library

of these direct selected fragments is created.

The cDNA clones isolated by direct selection are analyzed by two methods. Where the genomic target DNA sequence is obtained from a pool of BACs from the disorder region, the cDNAs are mapped to BAC genomic clones to verify their chromosomal location. This is accomplished by arraying the cDNAs in microtiter dishes, and replicating their DNA in high-density grids. Individual genomic clones known to map to the region are then hybridized to the grid to identify direct selected cDNAs mapping to that region. cDNA clones that are confirmed to correspond to individual BACs are sequenced. To determine whether the cDNA clones isolated by direct selection share sequence identity or similarity to previously identified genes, the DNA and protein coding sequences are compared to publicly available databases using the BLAST family of programs described above.

The combination of genomic DNA sequence and cDNA sequence provided by BAC sequencing and by direct cDNA selection yields an initial list of putative genes in the region. In the present invention, the genes in the region were candidates for the asthma locus. To further characterize each gene, Northern blots were performed to determine the size of the transcript corresponding to each gene, and to determine which putative exons were transcribed together to make an individual gene. For Northern blot analysis of each gene, probes are prepared from direct selected cDNA clones or by PCR amplifying specific fragments from genomic DNA, cDNA or from the BAC encoding the putative gene of interest. The Northern blot analysis is used to determine the size of the transcript and the tissues in which it is expressed. For transcripts that are not highly expressed, it is sometimes necessary to perform a reverse transcription PCR assay using RNA from the tissues of interest as a template for the reaction.

Gene identification by computational methods and by direct cDNA selection provides unique information about the genes in a region of a chromosome. Once genes are identified, it is possible to examine subjects for sequence variants. Variant sequences can be inherited as allelic differences

or can arise from spontaneous mutations.

Inherited alleles can be analyzed for linkage to a disease susceptibility locus. Linkage analysis is possible because of the nature of inheritance of chromosomes from parents to offspring. During meiosis, the two parental
5 homologs pair to guide their proper separation to daughter cells. While they are paired, the two homologs exchange pieces of the chromosomes, in an event called "crossing over" or "recombination." The resulting chromosomes contain parts that originate from both parental homologs. The closer together two sequences are on the chromosome, the less likely that a recombination
10 event will occur between them, and the more closely linked they are.

In the present invention, data obtained from the different families were combined and analyzed together by a computer using statistical methods described herein. The results were then used as evidence for linkage between the genetic markers used and an asthma susceptibility locus.

15 In general, a recombination frequency of 1% is equivalent to approximately 1 map unit, a relationship that holds up to frequencies of about 20% or 20 cM. One centimorgan (cM) is roughly equivalent to 1,000 kb of DNA. The entire human genome is 3,300 cM long. In order to find an unknown disease gene within 5-10 cM of a marker locus, the whole human
20 genome can be searched with roughly 330 informative marker loci spaced at approximately 10 cM intervals (Botstein et al., 1980, *Am. J. Hum. Genet.*, **32**:314-331).

The reliability of linkage results is established by using a number of statistical methods. The methods most commonly used for the detection by
25 linkage analysis of oligogenes involved in the etiology of a complex trait are non-parametric or model-free methods which have been implemented into the computer programs MAPMAKER/SIBS (L. Kruglyak and E.S. Lander, 1995, *Am. J. Hum. Genet.* **57**:439-454) and GENEHUNTER (L. Kruglyak et al., 1996, *Am. J. Hum. Genet.* **58**:1347-1363). Typically, linkage analysis is performed
30 by typing members of families with multiple affected individuals at a given marker locus and evaluating if the affected members (excluding parent-

offspring pairs) share alleles at the marker locus that are identical by descent (IBD) more often than expected by chance alone.

As a result of the rapid advances in mapping the human genome over the last few years, and concomitant improvements in computer methodology, it has become feasible to carry out linkage analyses using multi-point data. Multi-point analysis provides a simultaneous analysis of linkage between the trait and several linked genetic markers, when the recombination distance among the markers is known. A LOD score statistic is computed at multiple locations along a chromosome to measure the evidence that a susceptibility locus is located nearby. A LOD score is the logarithm base 10 of the ratio of the likelihood that a susceptibility locus exists at a given location to the likelihood that no susceptibility locus is located there. By convention, when testing a single marker, a total LOD score greater than +3.0 (that is, odds of linkage being 1,000 times greater than odds of no linkage) is considered to be significant evidence for linkage.

Multi-point analysis is advantageous for two reasons. First, the informativeness of the pedigrees is usually increased. Each pedigree has a certain amount of potential information, dependent on the number of parents heterozygous for the marker loci and the number of affected individuals in the family. However, few markers are sufficiently polymorphic as to be informative in all those individuals. If multiple markers are considered simultaneously, then the probability of an individual being heterozygous for at least one of the markers is greatly increased. Second, an indication of the position of the disease gene among the markers may be determined. This allows identification of flanking markers, and thus eventually allows identification of a small region in which the disease gene resides.

EXAMPLES

The examples as set forth herein are meant to exemplify the various aspects of the present invention and are not intended to limit the invention in any way.

5 **EXAMPLE 1: Family Collection**

Asthma is a complex disorder that is influenced by a variety of factors, including both genetic and environmental effects. Complex disorders are typically caused by multiple interacting genes, some contributing to disease development and some conferring a protective effect. The success of linkage
10 analyses in identifying chromosomes with significant LOD scores is achieved in part as a result of an experimental design tailored to the detection of susceptibility genes in complex diseases, even in the presence of epistasis and genetic heterogeneity. Also important are rigorous efforts in ascertaining asthmatic families that meet strict guidelines, and collecting accurate clinical
15 information.

Given the complex nature of the asthma phenotype, non-parametric affected sib pair analyses were used to analyze the genetic data. This approach does not require parameter specifications such as mode of inheritance, disease allele frequency, penetrance of the disorder, or phenocopy
20 rates. Instead, it determines whether the inheritance pattern of a chromosomal region is consistent with random segregation. If it is not, affected sibs inherit identical copies of alleles more often than expected by chance. Because no models for inheritance are assumed, allele-sharing methods tend to be more robust than parametric methods when analyzing complex disorders. They do,
25 however, require larger sample sizes to reach statistically significant results.

At the outset of the program, the goal was to collect 400 affected sib-pair families for the linkage analyses. Based on a genome scan with markers spaced ~10 cM apart, this number of families was predicted to provide > 95% power to detect an asthma susceptibility gene that caused an increased risk
30 to first-degree relatives of 3-fold or greater. The assumed relative risk of 3-fold was consistent with epidemiological studies in the literature that suggest an

increased risk ranging from 3- to 7-fold. The relative risk was based on gender, different classifications of the asthma phenotype (i.e. bronchial hyper-responsiveness versus physician's diagnosis) and, in the case of offspring, whether one or both parents were asthmatic.

- 5 The family collection efforts exceeded the initial goal of 400, obtaining a total of 444 affected sibling pair (ASP) families, with 342 families from the UK and 102 families from the US. The ASP families in the US collection were Caucasian with a minimum of two affected siblings that were identified through both private practice and community physicians as well as through advertising.
- 10 A total of 102 families were collected in Kansas, Nebraska, and Southern California. In the UK collection, Caucasian families with a minimum of two affected siblings were identified through physicians' registers in a region surrounding Southampton and including the Isle of Wight. In both the US and UK collections, additional affected and unaffected sibs were collected
- 15 whenever possible. An additional 39 families from the United Kingdom were utilized from an earlier collection effort with different ascertainment criteria. These families were recruited either: 1) without reference to asthma and atopy; or 2) by having at least one family member or at least two family members affected with asthma. The randomly ascertained samples were
- 20 identified from general practitioner registers in the Southampton area. For families with affected members, the probands were recruited from hospital based clinics in Southampton. Seven pedigrees extended beyond a single nuclear family.

- 25 Families were included in the study if they met all of the following criteria: 1) the biological mother and biological father were Caucasian and agreed to participate in the study; 2) at least two biological siblings were alive, each with a current physician diagnosis of asthma, and were 5 to 21 years of age; and 3) the two siblings were currently taking asthma medications on a regular basis. This included regular, intermittent use of inhaled or oral
- 30 bronchodilators and regular use of cromolyn, theophylline, or steroids.

Families were excluded from the study if they met any one of the

following criteria: 1) both parents were affected (i.e., with a current diagnosis of asthma, having asthma symptoms, or on asthma medications at the time of the study); 2) any of the siblings to be included in the study was less than 5 years of age; 3) any asthmatic family member to be included in the study was taking beta-blockers at the time of the study, 4) any family member to be included in the study had congenital or acquired pulmonary disease at birth (e.g. cystic fibrosis), a history of serious cardiac disease (myocardial infarction) or any history of serious pulmonary disease (e.g. emphysema); or 5) any family member to be included in the study was pregnant.

An extensive clinical instrument was designed and data from all participating family members were collected. The case report form (CRF) included questions on demographics, medical history including medications, a health survey on the incidence and frequency of asthma, wheeze, eczema, hay fever, nasal problems, smoking, and questions on home environment.

Data from a video questionnaire designed to show various examples of wheeze and asthmatic attacks were also included in the CRF. Clinical data, including skin prick tests to 8 common allergens, total and specific IgE levels, and bronchial hyper-responsiveness following a methacholine challenge, were also collected from all participating family members. All data were entered into a SAS dataset by IMTCI, a CRO; either by double data entry or scanning followed by on-screen visual validation. An extensive automated review of the data was performed on a routine basis and a full audit at the conclusion of the data entry was completed to verify the accuracy of the dataset.

EXAMPLE 2: Genome Scan

In order to identify chromosomal regions linked to asthma, the inheritance pattern of alleles from genetic markers spanning the genome was assessed on the collected family resources. As described above, combining these results with the segregation of the asthma phenotype in these families allows the identification of genetic markers that are tightly linked to asthma. In turn, this provides an indication of the location of genes predisposing affected individuals to asthma. The genotyping strategy was twofold: 1) to conduct a

genome wide scan using markers spaced at approximately 10 cM intervals; and 2) to target ten chromosomal regions for high density genetic mapping.

The initial candidate regions for high-density mapping were chosen based on suggestions of linkage to these regions by other investigators.

- 5 Genotypes of PCR amplified simple sequence microsatellite genetic linkage markers were determined using ABI model 377 Automated Sequencers (PE Applied Biosystems). Microsatellite markers were obtained from Research Genetics Inc. (Huntsville, AL) in the fluorescent dye-conjugated form (see Dubovsky et al., 1995, *Hum. Mol. Genet.* 4(3):449-452). The markers
- 10 comprised a variation of a human linkage mapping panel as released from the Cooperative Human Linkage Center (CHLC), also known as the Weber lab screening set version 8. The variation of the Weber 8 screening set consisted of 529 markers with an average spacing of 6.9 cM (autosomes only) and 7.0 cM (all chromosomes). Eighty-nine percent of the markers consisted of either
- 15 tri- or tetra-nucleotide microsatellites. There were no gaps present in chromosomal coverage greater than 17.5 cM.

- Study subject genomic DNA (5 µl; 4.5 ng/µl) was amplified in a 10 µl PCR reaction using AmpliTaq Gold DNA polymerase (0.225 U); 1 X PCR buffer (80 mM (NH₄)₂SO₄; 30 mM Tris-HCl (pH 8.8); 0.5% Tween-20); 200 µM each
- 20 dATP, dCTP, dGTP and dTTP; 1.5-3.5 µM MgCl₂; and 250 µM forward and reverse PCR primers. PCR reactions were set up in 192 well plates (Costar) using a Tecan Genesis 150 robotic workstation equipped with a refrigerated deck. PCR reactions were overlaid with 20 µl mineral oil, and thermocycled on an MJ Research Tetrad DNA Engine equipped with four 192 well heads using
- 25 the following conditions: 92°C for 3 min; 6 cycles of 92°C for 30 sec, 56°C for 1 min, 72°C for 45 sec; followed by 20 cycles of 92°C for 30 sec, 55°C for 1 min, 72°C for 45 sec; and a 6 min incubation at 72°C.

- PCR products of 8-12 microsatellite markers were subsequently pooled into two 96-well microtitre plates (2.0 µl PCR product from TET and FAM
- 30 labeled markers, 3.0 µl HEX labeled markers) using a Tecan Genesis 200 robotic workstation and brought to a final volume of 25 µl with H₂O. Following

this, 1.9 µl of pooled PCR product was transferred to a loading plate and combined with 3.0 µl loading buffer (2.5 µl formamide/blue dextran (9.0 mg/ml), 0.5 µl GS-500 TAMRA labeled size standard, ABI). Samples were denatured in the loading plate for 4 min at 95°C, placed on ice for 2 min, and
5 electrophoresed on a 5% denaturing polyacrylamide gel (FMC on the ABI 377XL). Samples (0.8 µl) were loaded onto the gel using an 8 channel Hamilton Syringe pipettor.

Each gel consisted of 62 study subjects and 2 control subjects (CEPH parents ID #1331-01 and 1331-02, Coriell Cell Repository, Camden, NJ).
10 Genotyping gels were scored in duplicate by investigators blind to patient identity and affection status using GENOTYPER analysis software V 1.1.12 (ABI; PE Applied Biosystems). Nuclear families were loaded onto the gel with the parents flanking the siblings to facilitate error detection. The final tables obtained from the GENOTYPER output for each gel analysed were imported
15 into a SYBASE Database.

Allele calling (binning) was performed using the SYBASE version of the ABAS software (Ghosh et al., 1997, *Genome Research* 7:165-178). Offsize bins were checked manually and incorrect calls were corrected or blanked. The binned alleles were then imported into the program MENDEL (Lange et al.,
20 1988, *Genetic Epidemiology*, 5:471) for inheritance checking using the USERM13 subroutine (Boehnke et al., 1991, *Am. J. Hum. Genet.* 48:22-25). Non-inheritance was investigated by examining the genotyping traces and, once all discrepancies were resolved, the subroutine USERM13 was used to estimate allele frequencies.

25 **EXAMPLE 3: Linkage Analysis**

Chromosomal regions harboring asthma susceptibility genes by linkage analysis of genotyping data and three separate phenotypes, asthma, bronchial hyper-responsiveness, and atopic status were identified as follows.

30 1. Asthma Phenotype: For the initial linkage analysis, the phenotype and asthma affection status were defined by a patient who answered the following questions in the affirmative: i) have you ever had

asthma; ii) do you have a current physician's diagnosis of asthma; and iii) are you currently taking asthma medications? Medications included inhaled or oral bronchodilators, cromolyn, theophylline, or steroids. Multipoint linkage analyses of allele sharing in affected individuals were performed using the

5 MAPMAKER/SIBS analysis program (L. Kruglyak and E.S. Lander, 1995, *Am. J. Hum. Genet.* **57**:439-454). The map location and distances between markers were obtained from the genetic maps published by the Marshfield medical research foundation (<http://www.marshmed.org/genetics/>). Ambiguous ordering of markers in the Marshfield map was resolved using the program

10 MULTIMAP (T.C. Matise et al., 1994, *Nature Genet.* **6**:384-390).^o

Using the discrete phenotype of asthma (yes/no), a candidate region was identified on chromosome 20 with a LOD score of 2.94, based on 462 nuclear families. Figure 1 displays the multipoint LOD score against the map location of the markers along chromosome 20. A Maximum LOD Score (MLS) of 2.94 was obtained at location 7.9 cM, 0.3 cM proximal to marker D20S906. A second MLS of 2.94 was obtained at marker D20S482 at location 12.1 cM. An excess sharing by descent (Identity By Descent (IBD) = 2) of 0.31 was observed at both maximum LOD scores. Table 2 lists the single and multipoint

15 LOD scores at each marker. Analyses were done using a conservative approach by weighting multiple sibling pairs within a sibship. When affected sib pairs were utilized in the linkage analyses without weighting the LOD score on chromosome 20 maximized at D20S482 with a value of 3.19. Thus, these data provided strong evidence for the presence of an asthma susceptibility

20 gene in this region of chromosome 20.

25

TABLE 2

Marker	Distance	Single-point	Multipoint
D20S502	0.5	0.7	2.4
D20S103	2.1	2.4	2.3
D20S117	2.8	1.2	2.0
GTC4ATG	6.3	2.4	2.5
GTC3CA	6.6	1.3	2.7
D20S906	7.6	2.9	2.9
D20S842	9.0	1.3	2.5
D20S181	9.5	1.8	2.6
D20S193	9.5	2.5	2.5

D20S889	11.2	1.6	2.6
D20S482	12.1	1.9	2.9
D20S849	14.0	0.8	2.0
D20S835	15.1	0.5	1.8
D20S448	18.8	1.4	1.4
D20S602	21.2	1.1	1.1
D20S851	24.7	1.0	0.8
D20S604	32.9	0.0	0.1
D20S470	39.3	0.0	0.1
D20S477	47.5	0.0	0.0
D20S478	54.1	0.0	0.0
D20S481	62.3	0.0	0.0
D20S480	79.9	0.0	0.0
D20S171	95.7	0.4	0.1

2. Phenotypic Subgroups: Nuclear families were ascertained by the presence of at least two affected siblings with a current physician's diagnosis of asthma, as well as the use of asthma medication. In the initial analysis (see above), the evidence was examined for linkage based on that dichotomous phenotype (asthma – yes/no). To further characterize the linkage signals, additional quantitative traits were measured in the clinical protocol. Since quantitative trait loci (QTL) analysis tools with correction for ascertainment was not available, the following approach was taken to refine the linkage and association analyses:

i. Phenotypic subgroups that could be indicative of an underlying genotypic heterogeneity were identified. Asthma subgroups were defined according to 1) bronchial hyper-responsiveness (BHR) to methacholine challenge; or 2) to atopic status using quantitative measures like total serum IgE and specific IgE to common allergens.

ii. Non-parametric linkage analyses were performed on subgroups to test for the presence of a more homogeneous sub-sample. If genetic heterogeneity was present in the sample, the amount of allele sharing among phenotypically similar siblings was expected to increase in the appropriate subgroup in comparison to the full sample. A narrower region of significant increased allele sharing was also expected to result unless the overall LOD score decreased as a consequence of having a smaller sample size and of using an approximate partitioning of the data.

iii. Alternatively, allele sharing probabilities were

parameterized as a function of the quantitative trait value of each child in a given sib pair, as advocated by N. Morton and implemented in his program BETA (N. Morton, 1996, *Proc. Natl. Acad. Sci. USA* **93**:3471-3476). This approach alleviated the need to dichotomize a quantitative trait. However, the program did not correct for the use of non-independent sib pairs in sibship of size 3 or larger. As such it did not provide an accurate measure of the significance of a linkage finding, but was used to corroborate the localization of the linkage signal.

3. Results for BHR and IgE: PC₂₀, the concentration of methacholine resulting in a 20% drop in FEV₁ (forced expiratory volume), was polychotomized in four groups and analyses were performed on the subsets of asthmatic children with mild to severe BHR (PC₂₀ ≤ 4 mg/ml) or PC₂₀(4), as well as on the broader subset with borderline to severe BHR (PC₂₀ ≤ 16 mg/ml) or PC₂₀(16). As shown in the LOD plot in Figure 2, the MLS for the subset of 127 nuclear families with at least two PC₂₀(4) affected sibs was 2.97 at 11.8 cM, 0.3 cM from D20S482, with an excess sharing by descent of 0.37. As shown in Figure 3, for the 218 nuclear families with at least two PC₂₀(16), the MLS was 3.93 at D20S482 with an excess sharing of 0.36. Both PC₂₀(4) and PC₂₀(16) strongly implicated the region of chromosome 20 under the second peak around marker D20S482. When considering the more extreme phenotype, PC₂₀(4), a higher proportion of families was linked to the region. However, the increase in LOD score for the PC₂₀(16) phenotype indicated that families concordant for the milder BHR phenotype also contributed to the linkage signal and would provide a larger pool of linked families.

Total IgE was dichotomized using an age specific cutoff for elevated levels (one standard deviation above the mean). Similarly, a dichotomous variable was created using specific IgE to common allergens. An individual was assigned a high specific IgE value if his/her level was positive (grass or tree) or elevated (> 0.35 KU/L for cat, dog, mite A, mite B, alternaria, or ragweed) for at least one such measure. In linkage analyses, the subset of asthmatic children with high total IgE (274 families) was given a maximum LOD

score of 2.3 at 11.6 cM (Figure 4), while the subset with high specific IgE (288 families) was given a LOD score of 1.87 at 12.1 cM (Figure 5). Similar to the BHR results, analyses based on IgE implicated the region under the second peak around marker D20S482. The substantially lower LOD scores using the subset of affected sibs concordant for atopy indicated the presence of groups with fewer linked families. Thus, atopy in asthmatic individuals was not the primary phenotype associated with the linkage signal on chromosome 20.

The BETA program (Morton, 1996) was used on two scales for PC₂₀. Individuals that did not drop 20% by the last dose administered (16 mg/ml) were assigned an arbitrary value of 32 mg/ml. First, a (0,1)-severity scale was constructed by applying a linear transformation to PC₂₀ where 0 mg/ml received a score of 1 and 32 mg/ml received a score of 0. For this scale, individuals that did not drop 20% in their FEV₁ did not contribute to the LOD score. A maximum LOD score of 3.43 was achieved at 12.1 cM with marker D20S482. Second, a linear transformation of PC₂₀ was used where 0 mg/ml received a score of 1 and 32 mg/ml a score of -1. In other words, in addition to the high concordant pairs, discordant pairs and concordant pairs that did not drop would also contribute to the LOD score. In contrast, individuals with PC₂₀ close to 16 mg/ml would have little impact on the LOD score. A maximum LOD score of 2.08 was again achieved at 12.1 cM.

Accordingly, a consistent pattern of evidence by linkage analysis pointed to the existence of an asthma susceptibility locus in the vicinity of marker D20S482. This was supported by the initial analysis of the asthma (yes/no) phenotype and by analyses of BHR in asthmatic individuals. Localization in the region of marker D20S482 was obtained using both BHR and IgE phenotypes.

EXAMPLE 4: Physical Mapping

The linkage results for chromosome 20 described above were used to delineate a candidate region for a disorder-associated gene located on chromosome 20. Gene discovery efforts were thus initiated in a 25 cM interval from the 20p telomere (marker D20S502) to marker D20S851, representing a >98% confidence interval. All genes known to map to this interval were

considered as candidates. Intensive physical mapping (BAC contig construction) focused on a 90% confidence interval between markers D20S103 and D20S916, a 15 cM interval. The discovery of novel genes using direct cDNA selection focused on a 95% confidence interval between markers
5 D20S502 (20p telomere) and D20S916, a 17 cM region.

The following section describes details of the efforts to generate cloned coverage of the disorder gene region on chromosome 20, i.e., construction of a BAC contig spanning the region. There were two primary reasons for using this approach: 1) to provide genomic clones for DNA sequencing (analysis of
10 this sequence would provide information about the gene content of the region); and 2) to provide reagents for direct cDNA selection (this would provide additional information about novel genes mapping to the interval). The physical map consisted of an ordered set of molecular landmarks, and a set of bacterial artificial chromosome clones (BACs; U.-J. Kim et al., 1996,
15 *Genomics* **34**:213-218; H. Shizuya et al., 1992, *Proc. Natl. Acad. Sci. USA* **89**:8794-8797) that contained the disorder gene region from human chromosome 20p13-p12.

Figure 6 depicts the BAC/STS content contig map of human chromosome 20p13-p12. Markers used to screen the RPCI-11 BAC library (P.
20 deJong, Roswell Park Cancer Institute (RPCI)) are shown in the top row. Markers that were present in the Genome Database (GDB, <http://gdbwww.gdb.org/>) are represented by GDB nomenclature. The BAC clones are shown below the markers as horizontal lines. BAC RPCI-11_1098L22 is labeled and the location of Gene 216, described herein, is
25 indicated at the top of the figure.

1. Map Integration. Various publicly available mapping resources were utilized to identify existing STS (sequence tagged site) markers (Olson et al., 1989, *Science*, **245**:1434-1435) in the 20p13-p12 region. Resources included the GDB (<http://gdbwww.gdb.org/>), Genethon ([http://www.genethon.
30 fr/genethon_en.html](http://www.genethon.fr/genethon_en.html)), Marshfield Center for Medical Genetics (<http://www.marshmed.org/genetics/>), the Whitehead Institute Genome Center

(<http://www-genome.wi.mit.edu/>), GeneMap98, dbSTS and dbEST (NCBI, <http://www.ncbi.nlm.nih.gov/>), the Sanger Centre (<http://www.sanger.ac.uk/>), and the Stanford Human Genome Center (<http://www-shgc.stanford.edu/>).

Maps were integrated manually to identify markers mapping to the disorder region. A list of the markers is provided in Table 3.

2. **Marker Development:** Sequences for existing STSs were obtained from the GDB, RHDB (<http://www.ebi.ac.uk/RHdb/>), or NCBI, and were used to pick primer pairs (overgos; see Table 3) for BAC library screening. Novel markers were developed either from publicly available genomic sequences, proprietary cDNA sequences, or from sequences derived from BAC insert ends (described below). Primers were chosen using a script that automatically performs vector and repetitive sequence masking using CROSSMATCH (P. Green, University of Washington). Subsequent primer selection was performed using a customized Filemaker Pro database (<http://www.filemaker.com>). Primers for use in PCR-based clone confirmation or radiation hybrid mapping (described below) were chosen using the program Primer3 (Steve Rozen, Helen J. Skaletsky, 1996, 1997, http://www-genome.wi.mit.edu/genome_software/other/primer3.html).

Table 3

Overgo	Locus	DNA Type	Gene	Forward Primer	SEQ ID NO	Reverse Primer	SEQ ID NO
stSG24277		Genomic		aacitcgaatgagaagcgtg	34	aaccaccacggattcaegcttc	45
stSG408		EST		aatatcatgcaccatgaccac	35	ataaccagatggctgtgggtca	46
A005005		EST	Attractin (ATTN)	tggagtgagttgttaacctat	36	atccccgcgatgaatagttta	47
B849D17AL		BACend		ggagcttatctctggattctta	37	gttgagagcccacttagataat	48
SN2		EST	Sialoadhesin (SN)	agagccacacatccatgctctg	38	gcatttgggggaagccagacat	49
AFM6026xh5	D20S867	MSAT		aagccactctgtgaattgccat	39	gccactagagggcaatggcaat	50
SN1		EST	Sialoadhesin (SN)	gagtagtctgtatgaccagatgg	40	cgaacgcacacagggccatctgg	51
stsH22126		EST		gtctggcgaatggagcatgaata	41	tcagggtcattcatttcatg	52
WI4876	D20S762	Genomic		attagagccatgaaggaaagg	42	tgcatacaacttctccttctct	53
stSG30448		EST		acaotgcttgggggacagct	43	agltgcagagacacctgctgtc	54
WI18677		EST		cacgagccacagagccagctc	44	ctgggagagagacggagctggc	55

20

3. **Radiation Hybrid (RH) Mapping:** Radiation hybrid mapping was performed against the Genebridge4 panel (Gyapay et al., 1996, *Hum. Mol. Genet.* 5:339-46) purchased from Research Genetics, in order to refine the chromosomal localization of genetic markers used in genotyping as well as to

identify, confirm, and refine localizations of markers from proprietary sequences. Standard PCR procedures were used for typing the RH panel with markers of interest. Briefly, 10 µl PCR reactions contained 25 ng DNA of each of the 93 Genebridge4 RH samples. PCR products were electrophoresed on 5 2% agarose gels (Sigma) containing 0.5 µg/ml ethidium bromide in 1 X TBE at 150 volts for 45 min. Model A3-1 electrophoresis systems were used (Owl Scientific Products, Portsmouth, NH). Typically, gels contained 10 tiers of lanes with 50 wells/tier. Molecular weight markers (100 bp ladder, GibcoBRL, Rockville, MD) were loaded at both ends of the gel. Images of the gels were 10 captured with a Kodak DC40 CCD camera and processed with Kodak 1D software (www.kodak.com). The gel data were exported as tab delimited text files; names of the files included information about the panel screened, the gel image files and the marker screened. These data were automatically imported using a customized Perl script into Filemaker databases for data storage and 15 analysis. The data were then automatically formatted and submitted to an internal server for linkage analysis to create a radiation hybrid map using RHMAPPER (L. Stein et al., 1995; available from Whitehead Institute/MIT Center for Genome Research, at <http://www.genome.wi.mit.edu/ftp/pub/software/rhmapper/>, and via anonymous ftp to <ftp://ftp.genome.wi.mit.edu>, 20 in the directory /pub/software/rhmapper.)

4. BAC Library Screening: The protocol used for BAC library screening was based on the "overgo" method, originally developed by John McPherson at Washington University in St. Louis (<http://www.tree.caltech.edu/protocols/overgo.html>, and W-W. Cai et al., 1998, *Genomics* **54**:387-397). 25 This method involved filling in the overhangs generated after annealing two primers, each 22 nucleotides in length, which overlap by 8 nucleotides. The resulting labeled 36 bp product was then used in hybridization-based screening of high density grids derived from the RPCI-11 BAC library (deJong, *supra*). Typically, 15 probes were pooled together to hybridize 12 filters (13.5 genome 30 equivalents).

Stock solutions (2 µM) of combined complementary oligos were heated

at 80°C for 5 min, placed at 37°C for 10 min, and then stored on ice. Labeling reactions included the following: 1.0 µl H₂O; 5 µl mixed oligos (2 µM each); 0.5 µl BSA (2 mg/ml); 2 µl OLB (-A, -C, -N6) Solution (see below); 0.5 µl ³²P-dATP (3000 Ci/mmol); 0.5 µl ³²P-dCTP (3000 Ci/mmol); and 0.5 µl Klenow fragment (5 U/µl). The reaction was incubated at room temperature for 1 hr, and unincorporated nucleotides were removed using Sephadex G50 spin columns.

Solution O: 1.25 M Tris-HCL, pH 8, 125 M MgCl₂; Solution A: 1 ml Solution O, 18 µl 2-mercaptoethanol, 5µl 0.1M dTTP, 5µl 0.1 M dGTP; Solution B: 2 M HEPES-NaOH, pH 6.6; Solution C: 3 mM Tris-HCl, pH 7.4, 0.2 mM EDTA; Solutions A, B, and C were combined to a final ratio of 1:2.5:1.5, and aliquots were stored at -20°C.

High-density BAC library membranes were pre-wetted in 2 X SSC at 58°C. Filters were then drained slightly and placed in hybridization solution (1% BSA; 1 mM EDTA, pH 8.0; 7% SDS; and 0.5 M sodium phosphate), pre-warmed to 58°C, and incubated at 58°C for 2-4 hr. Typically, 6 filters were hybridized in each container. Ten milliliters of pre-hybridization solution was removed, combined with the denatured overgo probes, and added back to the filters. Hybridization was performed overnight at 58°C. The hybridization solution was removed and filters were washed once in 2 X SSC, 0.1% SDS, followed by a 30 min wash in the same solution at 58°C. Filters were then washed in: 1) 1.5 X SSC and 0.1% SDS at 58°C for 30 min; 2) 0.5 X SSC and 0.1% SDS at 58°C for 30 min; and finally in 3) 0.1 X SSC and 0.1% SDS at 58°C for 30 min. Filters were then wrapped in Saran Wrap and exposed to film overnight. To remove bound probe, filters were treated in 0.1 X SSC and 0.1% SDS pre-warmed to 95°C and cooled room temperature. Clone addresses were determined as described by instructions supplied by RPCI.

To recover clonal BAC cultures from the library, a sample from the appropriate library well was plated by streaking onto LB agar (T. Maniatis et al., 1982, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) containing 12.5 µg/ml chloramphenicol (Sigma), and plates were incubated overnight. A single colony and a portion of the initial

streak quadrant were inoculated into 400 µl LB plus chloramphenicol in wells of a 96 well plate. Cultures were grown overnight at 37°C. For storage, 100 µl of 80% glycerol was added and the plates placed at -80°C.

To determine the marker content of clones, aliquots of the 96 well plate cultures were transferred to the surface of nylon filters (GeneScreen Plus, NEN) placed on LB/chloramphenicol Petri plates. Colonies were grown overnight at 37°C and colony lysis was performed by placing filters on pools of:
1) 10% SDS for 3 min; 2) 0.5 N NaOH and 1.5 M NaCl for 5 min; and 3) 0.5 M Tris-HCl, pH 7.5, and 1 M NaCl for 5 min. Filters were then air-dried and washed free of debris in 2 X SSC for 1 hr. The filters were air-dried for at least 1 hr and DNA was crosslinked linked to the membrane using standard conditions. Probe hybridization and filter washing were performed as described above for the primary library screening. Confirmed clones were stored in LB containing 15% glycerol.

In certain cases, polymerase chain reaction (PCR) was used to confirm the marker content of clones. PCR conditions for each primer pair were initially optimized with respect to MgCl₂ concentration. The standard buffer was 10 mM Tris-HCl (pH 8.3), 50 mM KCl, MgCl₂, 0.2 mM each dNTP, 0.2 µM each primer, 2.7 ng/µl human DNA, 0.25 units of AmpliTaq (Perkin Elmer) and MgCl₂ concentrations of 1.0 mM, 1.5 mM, 2.0 mM or 2.4 mM. Cycling conditions included an initial denaturation at 94°C for 2 min followed by 40 cycles at 94°C for 15 sec, 55°C for 25 sec, and 72°C for 25 sec followed by a final extension at 72°C for 3 min. Depending on the results from the initial round of optimization the conditions were further optimized if necessary. Variables included increasing the annealing temperature to 58°C or 60°C, increasing the cycle number to 42 and the annealing and extension times to 30 sec, and using AmpliTaqGold (Perkin Elmer).

5. BAC DNA Preparation: Several different types of DNA preparation methods were used for isolation of BAC DNA. The manual alkaline lysis miniprep protocol listed below (Maniatis et al., 1982) was successfully used for most applications, i.e., restriction mapping, CHEF gel

analysis and FISH mapping, but was not reproducibly successful in endsequencing. The Autogen protocol described below was used specifically for BAC DNA preparation for endsequencing.

For manual alkaline lysis BAC minipreps, bacteria were grown in 15 ml
5 terrific broth (TB) containing 12.5 µg/ml chloramphenicol. Cultures were placed
in a 50 ml conical tube at 37°C for 20 hr with shaking at 300 rpm. The cultures
were centrifuged in a Sorvall RT 6000 D at 3000 rpm (1800 x g) at 4°C for 15
min. The supernatant was then aspirated as completely as possible. In some
cases cell pellets were frozen at -20°C at this step for up to 2 weeks. The
10 pellet was then vortexed to homogenize the cells and minimize clumping.
Following this, 250 µl of P1 solution (50 mM glucose, 15 mM Tris-HCl, pH 8,
10 mM EDTA, and 100 µg/ml RNase A) was added and the mixture pipetted
up and down to mix. The mixture was then transferred to a 2 ml Eppendorf
tube. Subsequently, 350 µl of P2 solution (0.2 N NaOH, 1% SDS) was added,
15 mixed gently, and the mixture was incubated for 5 min at room temperature.
Then, 350 µl of P3 solution (3 M KOAc, pH 5.5) was added and mixed gently
until a white precipitate formed. The solution was incubated on ice for 5 min
and then centrifuged at 4°C in a microfuge for 10 min.

The supernatant was transferred carefully (avoiding the white
20 precipitate) to a fresh 2 ml Eppendorf tube, and 0.9 ml of isopropanol was
added; the solution was mixed and left on ice for 5 min. The samples were
centrifuged for 10 min, and the supernatant removed carefully. Pellets were
washed in 70% ethanol and air-dried for 5 min. Pellets were resuspended in
200 µl of TE8 (10 mM Tris-HCl, pH 8.0, 1.0 mM EDTA, pH 8.0), and RNase
25 (Boehringer Mannheim, <http://biochem.boehringer-mannheim.com>) added to
100 µg/ml. Samples were incubated at 37°C for 30 min, then precipitated by
addition of NH₄OAc to 0.5 M and 2 volumes of ethanol. Samples were
centrifuged for 10 min, and the pellets were washed with 70% ethanol. The
pellets were air-dried and dissolved in 50 µl TE8. Typical yields for this DNA
30 prep were 3-5 µg per 15 ml bacterial culture. Ten to 15 µl of DNA was used
for *EcoRI* restriction analysis; 5 µl was used for *NotI* digestion and clone insert

sizing by CHEF gel electrophoresis.

Autogen 740 BAC DNA preparations for endsequencing were made by dispensing 3 ml of LB media containing 12.5 µg/ml of chloramphenicol into autoclaved Autogen tubes. A single tube was used for each clone. For
5 inoculation, glycerol stocks were removed from -70°C storage and placed on dry ice. A small portion of the glycerol stock was removed from the original tube with a sterile toothpick and transferred into the Autogen tube. The toothpick was left in the Autogen tube for at least two min before discarding. After inoculation the tubes were covered with tape to ensure that the seal was
10 tight. When all samples were inoculated, the tubes were transferred into an Autogen rack holder and placed into a rotary shaker. Cultures were incubated at 37°C for 16-17 hr at 250 rpm. Following this, standard conditions for BAC DNA preparation, as defined by the manufacturer, were used to program the Autogen. However, samples were not dissolved in TE8 as part of the program.
15 DNA pellets were left dry.

When the program was completed, the tubes were removed from the output tray and 30 µl of sterile distilled and deionized H₂O was added directly to the bottom of the tube. The tubes were then gently shaken for 2-5 sec and then covered with parafilm and incubated at room temperature for 1-3 hr. DNA
20 samples were then transferred to an Eppendorf tube and used either directly for sequencing or stored at 4°C for later use.

6. BAC Clone Characterization: DNA samples prepared either by manual alkaline lysis or the Autogen protocol were digested with *EcoRI* for analysis of restriction fragment sizes. These data were used to compare the
25 extent of overlap among clones. Typically 1-2 µg were used for each reaction. Reaction mixtures included: 1 X Buffer 2 (NEB); 0.1 mg/ml BSA (NEB); 50 µg/ml RNase A (Boehringer Mannheim); and 20 units of *EcoRI* (NEB) in a final volume of 25 µl. Digestions were incubated at 37°C for 4-6 hr. BAC DNA was also digested with *NotI* for estimation of insert size by CHEF gel analysis (see
30 below). Reaction conditions were identical to those for *EcoRI*, except that 20 units of *NotI* were used. Six microliters of 6 X Ficoll loading buffer containing

bromphenol blue and xylene cyanol was added prior to electrophoresis.

EcoRI digests were analyzed on 0.6% agarose (Seakem, FMC Bioproducts, Rockland, ME) in 1X TBE containing 0.5 µg/ml ethidium bromide.

Gels (20 cm x 25 cm) were electrophoresed in a Model A4 electrophoresis unit (Owl Scientific) at 50 volts for 20-24 hr. Molecular weight size markers included undigested lambda DNA, *HindIII* digested lambda DNA, and *HaeIII* digested .X174 DNA. Molecular weight markers were heated at 65°C for 2 min prior to loading the gel. Images were captured with a Kodak DC40 CCD camera and analyzed with Kodak 1D software.

NotI digests were analyzed on a CHEF DRII (Bio-Rad) electrophoresis unit according to the manufacturer's recommendations. Briefly, 1% agarose gels (Bio-Rad pulsed field grade) were prepared in 0.5 X TBE, equilibrated for 30 min in the electrophoresis unit at 14 °C, and electrophoresed at 6 volts/cm for 14 hr with circulation. Switching times were ramped from 10 sec to 20 sec.

Gels were stained after electrophoresis in 0.5 µg/ml ethidium bromide. Molecular weight markers included undigested lambda DNA, *HindIII* digested lambda DNA, lambda ladder PFG ladder, and low range PFG marker (all from NEB).

7. BAC Endsequencing: The sequence of BAC insert ends utilized DNA prepared by either of the two methods described above. The ends of BAC clones were sequenced for the purpose of filling gaps in the physical map and for gene discovery information. The following vector primers specific to the BAC vector pBACe3.6 were used to generate endsequence from BAC clones: pBAC 5'-2 (TGT AGG ACT ATA TTG CTC; SEQ ID NO:56) and pBAC 3'-1 (CGA CAT TTA GGT GAC ACT; SEQ ID NO:57).

The ABI dye-terminator sequencing protocol was used to set up sequencing reactions for 96 clones. The BigDye (ABI; PE Applied Biosystems) Terminator Ready Reaction Mix with AmpliTaq[®] FS, Part number 4303151, was used for sequencing with fluorescently labeled dideoxy nucleotides. A master sequencing mix was prepared for each primer reaction set including: 1600 µl of BigDye terminator mix (ABI; PE Applied Biosystems); 800 µl of 5 X CSA

buffer (ABI; PE Applied Biosystems); 800 µl of primer (either pBAC 5'-2 or pBAC 3'-1 at 3.2 µM). The sequencing cocktail was vortexed to ensure it was well-mixed and 32 µl was aliquoted into each PCR tube. Eight microliters of the Autogen DNA for each clone was transferred from the DNA source plate to a corresponding well of the PCR plate. The PCR plates were sealed tightly and centrifuged briefly to collect all the reagents. Cycling conditions were as follows: 1) 95°C for 5 min; 2) 95°C for 30 sec; 3) 50°C for 20 sec; 4) 65°C for 4 min; 5) steps 2 through 4 were repeated 74 times; and 6) samples were stored at 4°C.

At the end of the sequencing reaction, the plates were removed from the thermocycler and centrifuged briefly. Centri-Sep 96 plates were then used according to manufacturer's recommendations to remove unincorporated nucleotides, salts, and excess primers. Each sample was resuspended in 1.5 µl of loading dye, and 1.3 µl of the mixture was loaded on ABI 377 Fluorescent Sequencers. The resulting endsequences were then used to develop markers to rescreen the BAC library for filling gaps and were also analyzed by BLASTN searching for EST or gene content.

EXAMPLE 5: Subcloning and Sequencing of BAC RPCI-11_1098L22

The physical map of the chromosome 20 region provided the location of the BAC RPCI-11_1098L22 clone that contains Gene 216 (see Figure 6). The BAC RPCI-11_1098L22 clone was deposited as clone RP11-1098L22 with the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209 USA, under ATCC Designation No. PTA-3171, on March 14, 2001 according to the terms of the Budapest Treaty. DNA sequencing of BAC, RPCI-11-1098L22 from the region was completed. BAC RPCI-11-1098L22 DNA, (the "BAC DNA") was isolated according to one of two protocols: either a QIAGEN purification (QIAGEN, Inc., Valencia, CA, per manufacturer's instructions) or a manual purification using a method which was a modification of the standard alkaline lysis/Cesium Chloride preparation of plasmid DNA (see e.g., F.M. Ausubel et al., 1997, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY). Briefly, for the manual

protocol, cells were pelleted, resuspended in GTE (50 mM glucose, 25 mM Tris-Cl (pH 8), 10 mM EDTA) and lysozyme (50 mg/ml solution), followed by addition of NaOH/SDS (1% SDS and 0.2N NaOH) and then an ice-cold solution of 3M KOAc (pH 4.5-4.8). RnaseA was added to the filtered
5 supernatant, followed by treatment with Proteinase K and 20% SDS. The DNA was then precipitated with isopropanol, dried, and resuspended in TE (10 mM Tris, 1 mM EDTA (pH 8.0)). The BAC DNA was further purified by cesium chloride density gradient centrifugation (Ausubel et al., 1997).

Following isolation, the BAC DNA was hydrodynamically sheared using
10 HPLC (Hengen et al., 1997, *Trends in Biochem. Sci.*, **22**:273-274) to an insert size of 2000-3000 bp. After shearing, the DNA was concentrated and separated on a standard 1% agarose gel. A single fraction, corresponding to the approximate size, was excised from the gel and purified by electroelution (Sambrook et al., 1989).

15 The purified DNA fragments were then blunt-ended using T4 DNA polymerase. The blunt-ended DNA was then ligated to unique *Bst*XI-linker adapters (5' GTCTTCACCACTGGGG (SEQ ID NO:58) and 5' GTGGTGAAGAC (SEQ ID NO:59) in 100-1000 fold molar excess). These linkers were complimentary to the *Bst*XI-cut pMPX vectors, while the overhang was not self-
20 complimentary. Therefore, the linkers would not concatemerize, nor would the cut-vector re-ligate to itself easily. The linker-adapted inserts were separated from unincorporated linkers on a 1% agarose gel and purified using GeneClean (BIO 101, Inc., Vista, CA). The linker-adapted insert was then ligated to a modified pBlueScript vector to construct a "shotgun" subclone library. The
25 vector contained an out-of-frame lacZ gene at the cloning site, which became in-frame in the event that an adapter-dimer was cloned. Such adapter-dimer clones gave rise to blue colonies, which were avoided.

All subsequent steps were based on sequencing by ABI377 automated DNA sequencing methods. Major modifications to the protocols are highlighted
30 below. Briefly, the library was transformed into DH5-competent cells (GibcoBRL, DH5-transformation protocol). Transformed cells were plated onto

antibiotic plates containing ampicillin and IPTG/X-gal. The plates were incubated overnight at 37°C. White colonies were identified and then used to plate individual clones for sequencing. The cultures were grown overnight at 37°C. DNA was purified using a silica bead DNA preparation method (Ng et al., 1996, *Nucl. Acids Res.*, **24**:5045-5047). In this manner, 25 µg of DNA was obtained per clone.

These purified DNA samples were sequenced using ABI dye-terminator chemistry. The ABI dye terminator sequence reads were run on ABI377 machines and the data were directly transferred to UNIX machines following lane tracking of the gels. All reads were assembled using PHRAP (P. Green, *Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V*, Jan. 1996, p.157) with default parameters and quality scores. The assembly was done at 8-fold coverage and yielded 1 contig, BAC RPCI-11-1098L22. SEQ ID NO:5 (Figure 7) comprises a portion of the BAC that includes the genomic sequence of Gene 216.

EXAMPLE 6: Gene Identification

Any gene or EST mapping to the interval based on public map data or proprietary map data was considered a candidate respiratory disease gene. Public map data were derived from several sources: the Genome Database (GDB, <http://gdbwww.gdb.org/>), the Whitehead Institute Genome Center (<http://www-genome.wi.mit.edu/>), GeneMap98, UniGene, OMIM, dbSTS and dbEST (NCBI, <http://www.ncbi.nlm.nih.gov/>), the Sanger Centre (<http://www.sanger.ac.uk/>), and the Stanford Human Genome Center (<http://www-shgc.stanford.edu/>). Proprietary data was obtained from sequencing genomic DNA (cloned into BACs) or cDNAs (identified by direct selection, screening of cDNA libraries or full length sequencing of IMAGE Consortium (<http://www-bio.11nl.gov/bbrp/image.html>) cDNA clones).

1. Gene Identification from clustered DNA fragments. DNA sequences corresponding to gene fragments in public databases (GenBank and human dbEST) and proprietary cDNA sequences (IMAGE consortium and direct selected cDNAs) were masked for repetitive sequences and clustered

using the PANGAEA Systems (Oakland, CA) EST clustering tool. The clustered sequences were then subjected to computational analysis to identify regions bearing similarity to known genes. This protocol included the following steps:

a. The clustered sequences were compared to the publicly available
5 UniGene database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997).
The parameters for this search were: $E = 0.05$, $v = 50$, $B = 50$, where E was
the expected probability score cutoff, V was the number of database entries
returned in the reporting of the results, and B was the number of sequence
alignments returned in the reporting of the results (Altschul et al., 1990).

10 b. The clustered sequences were compared to the GenBank database
(NCBI) using BLASTN2 (Altschul et al., 1997). The parameters for this search
were $E=0.05$, $V=50$, $B= 50$, where E , V , and B were defined as above.

c. The clustered sequences were translated into protein sequences for
all six reading frames, and the protein sequences were compared to a non-
15 redundant protein database compiled from GenPept Swissprot PIR (NCBI).
The parameters for this search were $E = 0.05$, $V = 50$, $B = 50$, where E , V , and
 B were defined as above.

d. The clustered sequences were compared to BAC sequences (see
below) using BLASTN2 (Altschul et al., 1997). The parameters for this search
20 were $E=0.05$, $V=50$, $B=50$, where E , V , and B were defined as above.

2. Gene Identification from BAC Genomic Sequence: Following
assembly of the BAC sequences into contigs, the contigs were subjected to
computational analyses to identify coding regions and regions bearing DNA
sequence similarity to known genes. This protocol included the following steps:

25 a. Contigs were degapped. The sequence contigs often contained
symbols (denoted by a period symbol) that represented locations where the
individual ABI sequence reads had insertions or deletions. Prior to automated
computational analysis of the contigs, the periods were removed. The original
data were maintained for future reference.

30 b. BAC vector sequences were "masked" within the sequence by using
the program crossmatch (P. Green, <http://chimera.biotech.washington>).

edu\UWGC). Since the shotgun library construction detailed above left some BAC vector in the shotgun libraries, this program was used to compare the sequence of the BAC contigs to the BAC vector and to mask any vector sequence prior to subsequent steps. Masked sequences were marked by "X" in the sequence files, and remained inert during subsequent analyses.

c. *E. coli* sequences contaminating the BAC sequences were masked by comparing the BAC contigs to the entire *E. coli* DNA sequence.

d. Repetitive elements known to be common in the human genome were masked using CROSSMATCH (P. Green, University of Washington). In this implementation of crossmatch, the BAC sequence was compared to a database of human repetitive elements (J. Jerka, Genetic Information Research Institute, Palo Alto, CA). The masked repeats were marked by "X" and remained inert during subsequent analyses.

e. The location of exons within the sequence was predicted using the MZEF computer program (Zhang, 1997, *Proc. Natl. Acad. Sci.*, **94**:565-568) and GenScan gene prediction program (Burge and Karlin, *J. Mol. Biol.*, **268**:78-94).

f. The sequence was compared to the publicly available UniGene database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were: $E = 0.05$, $v = 50$, $B = 50$, where E was the expected probability score cutoff, V was the number of database entries returned in the reporting of the results, and B was the number of sequence alignments returned in the reporting of the results (Altschul et al., 1990).

g. The sequence was translated into protein sequences for all six reading frames, and the protein sequences were compared to a non-redundant protein database compiled from GenPept Swissprot PIR (NCBI). The parameters for this search were $E = 0.05$, $V = 50$, $B = 50$, where E , V , and B were defined as above.

h. The BAC DNA sequence was compared to a database of clustered sequences using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V , and B were defined as above. The database of clustered sequences was prepared utilizing

a proprietary clustering technology (PANGEA Systems, Inc.) using cDNA clones derived from direct selection experiments (described below), human dbEST sequences mapping to the 20p13-p12 region, proprietary cDNAs, GenBank genes, and IMAGE consortium cDNA clones.

5 i. The BAC sequence was compared to the sequences derived from the ends of BACs from the region on chromosomes 20 using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V , and B were defined as above.

j. The BAC sequence was compared to the GenBank database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V , and B were defined as above.

k. The BAC sequence was compared to the STS division of GenBank database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V , and B were defined as above.

l. The BAC sequence was compared to the Expressed Sequence Tag (EST) GenBank database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V , and B were defined as above.

c. Mapping Analysis

Through mapping analysis, BAC RPCI-11_1098L22 (ATCC Designation No. PTA-3171) was identified as containing Gene 216. This BAC sequence (SEQ ID NO:5; Figure 7) included the genomic sequence of Gene 216 (SEQ ID NO:6; Figure 29), which corresponded to the cDNA sequence of Gene 216 (SEQ ID NO:1; Figure 24).

EXAMPLE 7: Gene 216 cDNA Cloning and Expression Analysis

1. Construction and screening of cDNA libraries: Directionally cloned cDNA libraries from normal lung and bronchial epithelium were constructed using standard methods (Soares et. al., 1994, *Automated DNA Sequencing and Analysis*, Adams et al. (eds), Academic Press, NY, pp. 110-

114). Total and cytoplasmic RNAs were extracted from tissue or cells by homogenizing the sample in the presence of Guanidinium Thiocyanate-Phenol-Chloroform extraction buffer (e.g. Chomczynski and Sacchi, 1987, *Anal. Biochem.*, **162**:156-159) using a polytron homogenizer (Brinkman Instruments, <http://www.brinkmann.com>). Poly A + RNA was isolated from total/cytoplasmic RNA using dynabeads-dT according to the manufacturer's recommendations (DynaI, Inc., <http://www.dynal.com>). The double stranded cDNA was then ligated into the plasmid vector pBluescript II KS+ (Stratagene, <http://www.stratagene.com>), and the ligation mixture was transformed into *E. coli* host DH10B or DH12S by electroporation (Soares, 1994). Following overnight growth at 37°C, DNA was recovered from the *E. coli* colonies after scraping the plates by processing as directed for the Mega-prep kit (QIAGEN). The quality of the cDNA libraries was estimated by counting a portion of the total number of primary transformants, determining the average insert size, and the percentage of plasmids with no cDNA insert. Additional cDNA libraries (human total brain, heart, kidney, leukocyte, and fetal brain) were purchased from Life Technologies (Bethesda, MD).

cDNA libraries, both oligo (dT) and random hexamer-primed, were used for isolating cDNA clones mapped within the disorder critical region. Four 10 x 10 arrays of each of the cDNA libraries were prepared as follows. The cDNA libraries were titrated to 2.5×10^6 using primary transformants. The appropriate volume of frozen stock was used to inoculate 2 L of LB/ampicillin (100 µg/µl). Four hundred aliquots containing 4 ml of the inoculated liquid culture were generated. Each tube contained about 5000 cfu (colony forming units). The tubes were incubated at 30°C overnight with shaking until an OD of 0.7-0.9 was obtained. Frozen stocks were prepared for each of the cultures by aliquotting 300 µl of culture and 100 µl of 80% glycerol. Stocks were frozen in a dry ice/ethanol bath and stored at -70°C. DNA was isolated from the remaining culture using the QIAGEN spin mini-prep kit according to the manufacturer's instructions. The DNA from the 400

cultures were pooled to make 80 column and row pools. Markers were designed to amplify putative exons from candidate genes. Once a standard PCR condition was identified and specific cDNA libraries were determined to contain cDNA clones of interest, the markers were used to screen the
5 arrayed library. Positive addresses indicating the presence of cDNA clones were confirmed by a second PCR using the same markers.

Once a cDNA library was identified as likely to contain cDNA clones corresponding to a transcript of interest from the disorder critical region, it was used to isolate a clone or clones containing cDNA inserts. This was
10 accomplished by a modification of the standard "colony screening" method (Sambrook et al., 1989). Specifically, twenty 150 mm LB plus ampicillin agar plates were spread with 20,000 cfu of cDNA library. Colonies were allowed to grow overnight at 37°C. Colonies were then transferred to nylon filters (Hybond from Amersham-Pharmacia, or equivalent) and duplicates prepared
15 by pressing two filters together essentially as described (Sambrook et al., 1989). The "master" plate was then incubated an additional 6-8 hr to allow the colonies additional growth. The DNA from the bacterial colonies was then bound to the nylon filters by treating the filters sequentially with denaturing solution (0.5 N NaOH, 1.5 M NaCl) for 2 min, and neutralization
20 solution (0.5 M Tris-Cl pH 8.0, 1.5 M NaCl) for 2 min (twice). The bacterial colonies were removed from the filters by washing in a solution of 2 X SSC/2% SDS for 1 min while rubbing with tissue paper. The filters were air-dried and baked under vacuum at 80°C for 1-2 hr to crosslink the DNA to the filters.

25 cDNA hybridization probes were prepared by random hexamer labeling (Fineberg and Vogelstein, 1983, *Anal. Biochem.*, **132**:6-13) or by including gene-specific primers and no random hexamers in the reaction (for small fragments). The colony membranes were then pre-washed in 10 mM Tris-Cl pH 8.0, 1 M NaCl, 1 mM EDTA, 0.1% SDS for 30 min at 55°C.
30 Following the pre-wash, the filters were pre-hybridized in > 2 ml/filter of 6 X SSC, 50 % deionized formamide, 2% SDS, 5 X Denhardt's solution, and 100

mg/ml denatured salmon sperm DNA, at 42°C for 30 min. The filters were then transferred to hybridization solution (6 X SSC, 2% SDS, 5 X Denhardt's, 100 mg/ml denatured salmon sperm DNA) containing denatured α -³²P-dCTP-labeled cDNA probe and incubated overnight at 42°C.

- 5 The following morning, the filters were washed under constant agitation in 2 X SSC, 2% SDS at room temperature for 20 min, followed by two washes at 65°C for 15 min each. A second wash was performed in 0.5 X SSC, 0.5% SDS for 15 min at 65°C. Filters were then wrapped in plastic wrap and exposed to radiographic film. Individual colonies on plates were
- 10 aligned with the autoradiograph and positive clones picked into a 1 ml solution of LB Broth containing ampicillin. After shaking at 37°C for 1-2 hr, aliquots of the solution were plated on 150 mm plates for secondary screening. Secondary screening was identical to primary screening (above) except that it was performed on plates containing ~250 colonies so that
- 15 individual colonies could be clearly identified. Positive cDNA clones were characterized by restriction endonuclease cleavage, PCR, and direct sequencing to confirm the sequence identity between the original probe and the isolated clone.

- 20 To obtain the full-length cDNA, novel sequence from the 5'-end of the clone was used to reprobe the library. This process was repeated until the length of the cDNA cloned matched that of the mRNA, estimated by Northern analysis. Utilizing this process, a single uterus clone was isolated and deposited as clone Gene 216_CS759 with the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209 USA,
- 25 under ATCC Designation No. PTA-3173, on March 14, 2001, according to the terms of the Budapest Treaty. The uterus clone (SEQ ID NO:3) contained the entire Gene 216 open reading frame. Both strands of this clone were completely sequenced and the data were compared against the BAC sequence. Any discrepancies were flagged, and these regions were
- 30 resequenced. The final analysis of the sequence revealed that the uterine clone was 3433 bp long and contained the full complement of exons defining

the open reading frame (SEQ ID NO:3). In addition, the clone contained a small portion of the 5' untranslated region (5 bp), the entire 3' untranslated region with a polyadenylation signal, and a poly A tail of 76 bp in length. The Gene 216 open reading frame was determined to be 2436 bp in length and to encode a protein of 812 amino acids (SEQ ID NO:363). Analysis of the composition of SNPs across the cDNA clone revealed that it contained the most frequent haplotype (Figure 8, see below).

Rapid Amplification of cDNA ends (RACE) was performed following the manufacturer's instructions using a Marathon cDNA Amplification Kit (CLONTECH) as a method for cloning the 5' and 3' ends of candidate genes. cDNA pools were prepared from total RNA by performing first strand synthesis. For first strand synthesis, a sample of total RNA sample was mixed with a modified oligo (dT) primer, heated to 70°C, cooled on ice and incubated with: 5 X first strand buffer (CLONTECH), 10 mM dNTP mix, and AMV Reverse Transcriptase (20 U/μl). The reaction mixture was incubated at 42°C for 1 hr and placed on ice. For second-strand synthesis, the following components were added directly to the reaction tube: 5 X second-strand buffer (CLONTECH), 10 mM dNTP mix, sterile water, and 20 X second-strand enzyme cocktail (CLONTECH). The reaction mixture was incubated at 16°C for 1.5 hr. T4 DNA Polymerase was added to the reaction mixture and incubated at 16°C for 45 min. The second-strand synthesis was terminated with the addition of an EDTA/Glycogen mix. The sample was purified by phenol/chloroform extraction and ammonium acetate precipitation. The cDNA pools were checked for quality by analyzing on an agarose gel for size distribution. Marathon cDNA adapters were then ligated onto the cDNA ends. The specific adapters contained priming sites that allowed for amplification of either 5' or 3' ends, and varied depending on the orientation of the gene specific primer (GSP) that was chosen. An aliquot of the double stranded cDNA was added to the following reagents: 10 μM Marathon cDNA adapter, 5 X DNA ligation buffer, T4 DNA ligase. The reaction was incubated at 16°C overnight and heat inactivated to terminate

the reaction. PCR was performed by the addition of the following to the diluted double stranded cDNA pool: 10X cDNA PCR reaction buffer, 10 μ M dNTP mix, 10 μ M GSP, 10 μ M AP1 primer (kit), 50 X Advantage cDNA Polymerase Mix. Thermal Cycling conditions were carried out at 94°C for 30 sec; 5 cycles of 94°C for 5 sec, 72°C for 4 min, 5 cycles of 94°C for 5 sec, and 70°C for 4 min; 23 cycles of 94°C for 5 sec; 68°C for 4 min. The first round of PCR was performed using the GSP to extend to the end of the adapter to create the adapter primer-binding site. Following this, exponential amplification of the specific cDNA of interest was performed.

Usually, a second, nested PCR was performed to provide specificity. The RACE product was analyzed on an agarose gel. Following excision from the gel and purification (GeneClean, BIO 101), the RACE product was then cloned into pCTNR (General Contractor DNA Cloning System, 5' - 3', Inc.) and sequenced to verify that the clone was specific to the gene of interest.

The 5' RACE technique was employed to identify the 5' untranslated region of Gene 216. Experiments were performed using lung mRNA and a primer that hybridized near the 5' end of the available sequence. The result of the experiment identified an additional 75 bp 5' of that present in the uterus cDNA clone (rt690; SEQ ID NO:351). This sequence was subsequently cloned and deposited with the ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209 USA), as clone Gene 216_rt690, under ATCC Designation No.PTA-3172 on March 14, 2001, according to the terms of the Budapest Treaty.

Further attempts to extend the 5' end of Gene 216 by 5' RACE gave similar results indicating that the 5' end of the transcript was obtained.

This sequence in combination with the uterus cDNA clone yielded the master consensus sequence containing the 5' to 3' cDNA for Gene 216 (SEQ ID NO:1; Figure 24).

2. Identification of Splice Variants: Additional cDNA clones were isolated that represented alternatively spliced variants of Gene 216. To ensure that all splice variants present in lung tissue were identified, an RT-PCR-based

screening protocol was designed using multiple primer pairs spanning the entire gene. These amplicons produced PCR fragments of approximately 600 bp and overlapped by approximately 100 bp. The PCR products were fractionated on agarose gels and any fragments that were different from the expected size were cloned and sequenced. These results are summarized in Figures 9 and 10. The availability of the complete genomic sequence of BAC RPCI-11_1098L22 enabled the intron/exon structure of Gene 216 (Figure 11) to be determined. Gene 216 contains 21 exons that span approximately 23.5 kb of genomic DNA.

Analysis of the sequence surrounding the intron/exon boundaries indicated that the consensus splice sequence GT/AG was upheld in all cases (Table 4). However, in several of the cDNA clones, an alternative use of a splice site at the intron/exon boundary of exon T was identified. The sequence CAGCAG was present at the border of intron ST and exon T resulting in a duplication of the canonical acceptor splice consensus CAG. Typically, a C residue preceding the AG is found in approximately 65% of acceptor splice sites. As a consequence, the splicing machinery can utilize either AG resulting in the presence or absence of an alanine. If the first AG (splice site 1) were utilized near the junction of intron ST and exon T, the resulting protein would encode the amino acid sequence DPQADQVQM (Figure 12) (SEQ ID NO:60). However, if the second AG (splice site 2) were favored, then one alanine would be omitted from the amino acid sequence and the protein would contain the amino acid sequence DPQDQVQM (Figure 12) (SEQ ID NO:61). The percentage that used splice site 1 or splice site 2 could not be determined from the dataset because the majority of the clones were derived from PCR-based techniques.

TABLE 4

EXON	3' INTRON	5' EXON	3' EXON	5' INTRON
A			AAG	GTGAGG
B	CAG	GAC	CCG	GTCAAGT
C	CAG	GTC	CCA	GTGAGT
D	CAG	CAG	ACG	GTGAGA
D (ALT)	CAG	CAG	GAG	GTACCC
E	TAG	GAT	GAG	GTGAGC
F	TAG	TGG	AGG	GTGAGG
G	CAG	GGC	CTG	GTGAGG
H	CAG	TTC	CAG	GTGGGG
I	CAG	CTT	CAC	GTGGGT
J	CAG	GGG	ACG	GTGAGC
K	CAG	GAC	CCG	GTACGC
L	TAG	GCA	CAG	GTTAAG
L2	CAG	GAG	CTG	GTGAGG
M	CAG	CTG	CTG	GTGAGA
N	CAG	GCT	GAG	GTAGGG
O	CAG	GGA	ATG	GTGAGC
O (ALT)	TAG	ATG	ATG	GTGAGC
U	TAG	GTG	GGG	GTGAGA
P	CAG	GTT	AAA	GTATGC
Q	CAG	ACC	TGG	GTAGGC
R	CAG	CCC	TGG	GTGAGT
S	CAG	ACC	AAG	GTAGGC
T	CAG	CAG		

C₆₅ A₁₀₀ G₁₀₀ N A₆₄ G₇₃ G₁₀₀ T₁₀₀ A₆₂ A₆₈ G₈₄ T₆₃

3. **Promoter Analysis:** In order to identify the transcriptional start site of Gene 216, multiple 5' RACE products were sequenced from several different tissues. In most cases the 5' ends were located 80 bp upstream of the translational start site. The region upstream of this sequence was then analyzed for potential transcription factor binding sites using GEMS Launcher, a promoter analysis program (<http://anthea.gsf.de/>). GEMS Launcher uses statistically weighted algorithms to identify binding elements that comprise a promoter or regulatory module. A stretch of DNA sequence spanning the 2000 bp upstream of the translational start site was analyzed. The results indicated that Gene 216 did not possess a TATA or CCAAT box. In fact, the first binding element that was identified was a GC box within the 5' untranslated region oriented in the opposite direction (Figure 13). This result is not unprecedented since 60% of TATA-less genes possess a GC box on the opposing strand. Also, this result was in agreement with published data regarding the promoters of mouse ADAM 17 and 19. Other binding elements that were identified within

600 bp upstream of the initiator methionine included an E-box, one AP2, and three SP1 sites (Figure 13). These types of binding elements were also identified in the mouse ADAM 17 and 19 genes and may represent components of a promoter module for Gene 216. Approximately 1200 bp upstream of the putative promoter module, GEMS Launcher identified binding elements that may comprise an additional regulatory element (Figure 13). This region was highly conserved with the mouse ortholog of Gene 216 (see below), as determined by dot matrix analysis.

4. BLAST Analysis: BLASTP, BLASTN, and BLASTX analysis of Gene 216 against protein and nucleotide databases revealed that it was a novel member of the ADAM (A Disintegrin And Metalloprotease) gene family. This gene family, of which there are currently 31 members, is a sub-group of the zinc-dependent metalloprotease superfamily. ADAMs have a complex domain organization that includes a signal sequence, propeptide, metalloprotease, disintegrin, cysteine-rich, and epidermal growth factor-like domains, as well as a transmembrane region and cytoplasmic tail. ADAM proteins have been implicated in many processes such as proteolysis in the secretory pathway and extracellular matrix, extra- and intra-cellular signaling, processing of plasma membrane proteins and procytokine conversion. The homology of Gene 216 and human ADAMs 19, 12, 15, 8 and 9 indicated that Gene 216 belonged to a branch of the 31-member family containing active metalloprotease domains (Figure 14).

6. Expression Analysis: To characterize the expression of Gene 216, a series of expression experiments were performed.

i. Northern Analysis: To characterize novel genes, Northern analysis (Sambrook et al., 1989) can be used to determine the length, in nucleotides, of the processed transcript or messenger RNA (mRNA). Probes were generated using one of the methods described below. Briefly, sequence verified IMAGE consortium cDNA clones were digested with appropriate restriction endonucleases to release the insert. The restriction digest was electrophoresed on an agarose gel and the bands containing the insert were

excised. The gel piece containing the DNA insert was placed in a Spin-X (Corning Costar Corporation, Cambridge, MA) or Supelco spin column (Supelco Park, PA) and spun at high speed for 15 min. The DNA was ethanol precipitated and resuspended in TE. Alternatively, PCR products obtained from genomic DNA or RT-PCR were purified. First, oligonucleotide primers were designed for use in the polymerase chain reaction (PCR) so that portions of the cDNA, EST, or genomic DNA could be amplified from a pool of DNA molecules or RNA population (RT-PCR). The PCR primers were used in a reaction containing genomic DNA to verify that they generated a product of the predicted size (based on the genomic sequence. Inserts purified from IMAGE clones or PCR products were random primer labeled (Fineberg and Vogelstein, *supra*) to generate probes for hybridization. Probes from purified PCR products were generated by incorporation of α -³²P-dCTP in second round of PCR. Commercially available Multiple Tissue Northern blots (CLONTECH) were hybridized and washed under conditions recommended by the manufacturer. A separate filter that contained 6 tissues from the immune system was also utilized. The results revealed a major 5.0 kb transcript and a minor 3.5 kb transcript that were expressed in most tissues examined (Figures 15A-15B). The strongest signals were consistently identified in heart, skeletal muscle, colon, lymph, and small intestine, with lung, liver, kidney, placenta, bone marrow, and brain showing moderate expression levels.

The 5 kb transcript was further analyzed to determine if it was an incompletely spliced version of the Gene 216 transcript. To test this hypothesis, Northern blotting was performed using cytoplasmic mRNA isolated from bronchial smooth muscle cells. The same radioactive probe was employed as previously. The results showed a very strong 3.5 kb signal and no signal at 5.0 kb (Figure 15C) suggesting that the predominant 5 kb transcript contained intronic material and was localized to the nucleus. Interestingly, intron QR is 1.4 kb in size. The addition of the QR intron and the 3.5 kb full length cDNA would total ~5.0 kb. Accordingly, there may be regulatory elements within the region around intron QR that affect splicing,

retention in the nucleus, and/or transport to the cytoplasm.

ii. RNA Dot Blot Analysis: RNA dot blotting was used to determine the expression of Gene 216 in a wide range of tissues. mRNA from 50 tissues was dotted onto a nylon filter, and a radioactive probe
5 designed to hybridize to the 3' untranslated region was used. Figure 16 shows that Gene 216 was highly expressed in gastrointestinal tissues as well as aorta, uterus, prostate, ovary, lung, fetal lung, trachea and placenta.

Notably, the majority of these tissues are derived from the endoderm, which forms a tube that produces the primordium of the digestive tract. Extensions
10 from this wall also develop into organs such as the lung and trachea.

iii. RT-PCR: Total RNA isolated from primary cultures of seven cell types cultured from lung tissue was analyzed in RT-PCR experiments. Genomic DNA was removed from the total RNA by DNaseI digestion. The "Superscript" Preamplification System for First strand cDNA synthesis" (Life
15 Technologies) was used according to manufacturer's specifications with oligo(dT) or random hexamers to synthesize cDNA from the DNaseI treated total RNA. Gene specific primers were used to amplify the target cDNAs in a 30 µl PCR reaction containing 0.5 µl of first strand cDNA, 1 µl sense primer (10 µM), 1 µl antisense primer (10 µM), 3 µl dNTPs (2 mM), 1.2 µl MgCl₂ (25 mM),
20 3 µl 10 X PCR buffer and 1 unit of Taq Polymerase (Perkin Elmer). The PCR reaction was initially incubated at 94°C for 4 min, followed by 30 cycles of incubation at 94°C for 30 sec, 58°C for 1 min, and 72°C for 1 min; then followed by a final incubation at 72°C for 7 min. PCR products were analyzed on agarose gels. Figure 17 shows that Gene 216 was expressed in lung
25 fibroblasts, pulmonary artery smooth muscle cells, bronchial smooth muscle cells and total lung, but not in bronchial epithelium or pulmonary artery endothelial cells.

iv. cDNA Library Representation: A comprehensive approach to determining the tissue distribution of Gene 216 was performed *in silico* by
30 mining the public EST database and Genome Therapeutics Corporation's internal cDNA database. BLAST analysis identified ESTs from multiple cDNA

libraries. A summary of all tissues expressing Gene 216 is given in Table 5.

TABLE 5

Source	Tissue
UNIGENE	Eye Muscle Placenta Stomach Uterus Whole embryo Breast Normal testis
Direct selected cDNAs	Bronchial smooth muscle (1 clone) Normal lung (2 clones) Brain (1 clone)
Primary cell types (RT/PCR)	Pulmonary artery smooth muscle Bronchial smooth muscle Lung fibroblast Total lung
RNA Dot Biot	Aorta Colon Bladder Uterus Prostate Ovary Small intestine Heart Stomach Testis Appendix Lung Trachea Fetal kidney Fetal lung
Northern Blot	Brain Heart Skeletal muscle Colon Thymus Spleen Kidney Liver Small intestine Placenta Lung Lymph Bone marrow

EXAMPLE 8: Gene 216 Polypeptide

1. ADAM Family Features: The zinc-dependent metalloprotease superfamily is comprised of several sub-groups. Those proteases that exhibit the characteristic Zn-binding consensus sequence HEXXHXXGXXH (SEQ ID NO:62) are referred to as zincins. The 3 histidines play an essential role in binding to the catalytically essential zinc ion. The zincins can be further classified into metzincins if a methionine residue is located beneath the active-site zinc ion ("Met-turn" motif). Within this sub-group there are 4 sub-families: astacins, matraxins, adamlysins, and serralsins. The ADAM genes fall within the adamlysins sub-family along with snake venom metalloproteases.

Currently, there are 31 members of the ADAM family. The ADAM genes encode proteins of approximately 750 amino acids with 8 different domains. Domain I is a pre-domain and contains the signal sequence peptide that facilitates secretion through the plasma membrane. Domain II is a pro-domain that is cleaved before the protein is secreted resulting in activation of the catalytic domain. Domain III is a catalytic domain containing metalloprotease activity. Domain IV is a disintegrin-like domain and is believed to interact with integrins or other receptors. Domain V is a cysteine-rich domain and is speculated to be involved in protein-protein interactions or in the presentation of the disintegrin-like domain. Domain VI is an EGF-like domain that plays a role in stimulating membrane fusion. Domain VII is a transmembrane domain that anchors the ADAM protein to the membrane. Domain VIII is a cytoplasmic domain and contains binding sites for cytoskeletal-associated proteins and/or SH3 binding domains that may play a role in bi-directional signaling. See Figure 8 for the location of ADAM domains identified in the Gene 216 protein sequence.

To determine whether Gene 216 was a novel member of the ADAM family, the 812 amino acid sequence was aligned by Pile-Up (Genetics Computer Group, <http://www.gcg.com>) (Figure 18). These analyses indicated that Gene 216 possessed the characteristic consensus sequence HEXXHXXGXXH (SEQ ID NO:62) located within the catalytic domain. In

addition, a methionine residue referred to as a "Met-turn" was identified in the Gene 216 protein. A conserved cysteine (amino acid 133 in Gene 216) that plays a role in activating ADAM proteins was identified in the prodomain of Gene 216 protein. In ADAM proteins, this single cysteine residue forms an intramolecular complex with the zinc ion bound to the metalloprotease domain and blocks the active site. The catalytic domain is activated by the dissociation of the cysteine from the complex, resulting in either a conformational change or enzymatic cleavage of the prodomain. This process is referred to as the "cysteine switch".

In ADAM 12, the position of the cysteine residue was reported to be located in a different position in the prodomain (B.L. Gilpin et al., 1998, *J. Biol. Chem.* **273**:157-166). This location would correspond to the cysteine residue at amino acid 179 in Gene 216 (Figure 19). However, in accordance with analyses performed by Stone et al., using 14 ADAMs, including ADAMs 8, 9, 12 and 15, the cysteine residue corresponding to position 133 of Gene 216 (Figures 18 and 19) was identified as being involved in the "cysteine switch". In addition, there appeared to be more sequence identity around the cysteine at amino acid 133 in Gene 216 than at position 179. This provided further support that the cysteine at position 133 was involved in the "cysteine switch". The alignment also indicated that the amino acid sequence of Gene 216 contained all eight domains that define the hallmarks of these types of genes (Figure 18).

Hydrophobicity analysis (PepPlot, Genetics Computer Group) of the Gene 216 amino acid sequence revealed the presence of two hydrophobic regions (Figure 20). One region is located at the amino terminus of the protein and is the putative the signal sequence. The other hydrophobic region is located near the carboxyl terminus and is the putative transmembrane domain that anchors the protein to the cell surface. Computational biology analysis (<http://blocks.fhcrc.org>) of the Gene 216 cytoplasmic domain revealed the presence of a putative SH2 and SH3 binding domain as well as a putative

casein kinase I phosphorylation site (Figure 19). These sites may contribute to a role in bi-directional signaling, a function attributed to ADAM proteins.

Sequence analyses indicated that Gene 216 is a novel member of the ADAM family. Gene 216 is most closely related to ADAMs 8, 9, 12, 15, and 19, a branch of the family that is known to possess an active metalloprotease domain. Table 6 lists the 5 most similar BLASTP hits using the Gene 216 amino acid sequence as a query. Based on BLASTN and BLASTP analysis, Gene 216 nucleotide sequence shares the 37% identity with the ADAM 19 nucleotide sequence; and Gene 216 amino acid sequence shares 58% identity with the ADAM 19 amino acid sequence.

Table 6: Top 5 Hits from BLAST Analysis of Gene 216 protein

Hit	GenBank Locus	Description	Smallest Sum
1	U66003	<i>Xenopus laevis</i> (ADAM 13)	5.5e-166
2	AF019887	<i>Mus musculus</i> metalloprotease-disintegrin meltrin beta	1.2e-139
3	AF134707	<i>Homo sapiens</i> disintegrin and metalloprotease domain 19 (ADAM19)	1.6e-139
4	S60257	Mouse mRNA for meltrin alpha	1.8e-121
5	AF023476	<i>Homo sapiens</i> meltrin-L precursor (ADAM12)	4.9e-119

Table 7 lists the top two hits from BLIMPS analysis of the Block protein motif database (<http://blocks.fhcrc.org/>).

Table 7: Top 2 Hits from BLIMPS Analysis of Gene 216 protein

Description	Strength Score	AA#	AA	Sequence
Disintegrins proteins	1950	1597	377	CCfAhnCsLRPGAQCAh- GdCCvRCIIKpAGaI- CRqAMGDCCIPEfCT- GTSshCPP (SEQ ID NO:335)
Zinc metallopeptidases	1173	1276	276	TMAHEIGHSLG (SEQ ID NO:336)

2. Amino Acid Changes: In total, there were 9 SNPs within the open reading frame of Gene 216. See Example 10 for details on polymorphism identification and Figure 19 for resulting changes to the protein sequence.

- 5 Seven of the nine SNPs constituted an amino acid change and the other 2 were synonymous. Of the 7 amino acid changes, 4 were clustered toward the carboxyl terminus of the protein: one within the identified transmembrane domain and 3 within the identified cytoplasmic domain.

- One SNP located in an identified SH2 binding domain resulted in a
10 significant amino acid change: methionine (hydrophobic) to threonine (polar). The remaining two SNPs in the identified cytoplasmic domain resulted in significant amino acid changes: proline (hydrophobic) to serine (polar) and glutamine (polar) to histidine (basic). These amino acid changes may disturb the signaling properties of the Gene 216 protein. In addition, the valine to
15 isoleucine amino acid change in the putative transmembrane domain may affect signaling efficiency.

- The two SNPs in the identified pro-domain generated significant amino acid changes: tyrosine (polar) to histidine (basic) and threonine (polar) to alanine (hydrophobic). Since the ADAM pro-domain is cleaved during
20 activation of the catalytic domain, it is possible that these amino acid changes affect the cleavage process. One SNP in the identified catalytic domain resulted in a change from alanine (hydrophobic) to valine (hydrophobic). This amino acid change may affect sheddase efficiency.

- Notably, amino acid changes in the identified Gene 216 catalytic
25 domain, especially within the metalloprotease domain, would be of great interest, as this domain is critical to sheddase function. Recently, the X-ray crystallographic data of the snake venom catalytic domain was determined and deposited in the public domain (<http://www.rcsb.org/pdb/cgi/explore.cgi?pid=9267984771616&pdbld=1C9G>; Accession No. 1C9GA). This information
30 can be utilized to determine whether an amino acid change alters the folding of the catalytic domain of the Gene 216 protein. In particular, the sequence of

the catalytic domain of Gene 216 protein can be plotted as X-ray crystallographic coordinates and used to determine changes in the tertiary structure of this domain.

3. Biological Role of Gene 216: ADAMs are part of a very large
5 superfamily called zinc-dependent metalloproteases (Stone et. al., 1999, *J. Prot. Chem.* **18**:447-465). Gene 216 represents a novel member of the ADAM family that is closely related to ADAM 19, a gene that was found to participate in the proteolytic processing of the membrane anchored protein neuregulin 1 (NRG1) (Shirakabe et. al., 2001, *J. Biol. Chem.* **276**(12):9352-8). The
10 expression and activation of ADAM 19 protein has been localized to the trans-Golgi apparatus. This has been observed for other ADAM proteins (Lum et al., 1998, *J. Biol. Chem.* **273**:26236-26247; Roghani et. al., 1999, *J. Biol. Chem.* **274**:3531-3540; Shirakabe et. al., 2001, *J. Biol. Chem.* **276**(12):9352-8). These data suggest that the ADAM genes, and Gene 216, encode proteins
15 that function in the trans-Golgi apparatus as intracellular processing enzymes. The processed substrates of these enzymes may be released into the cytosol as part of a signal transduction cascade leading to the cell surface.

The substrate of ADAM 19, NRG1, belongs to a group of growth factors (neuregulins) that are members of the epidermal growth factor family. The
20 neuregulins participate in an array of biological effects that are mediated by the epidermal growth factor family of tyrosine kinase receptors. Data suggest that the proteolytically cleaved isoform of NRG1, NRG- β 1, may induce the tyrosine phosphorylation of EGFR2 and EGFR3 in differentiated muscle cells (Shirakabe et. al., 2001, *J. Biol. Chem.* **276**(12):9352-8). The sequence
25 similarity of Gene 216 protein and ADAM 19 protein suggests that the neuregulins or their isoforms serve as substrates for Gene 216 protein. The Gene 216-processed neuregulins or isoforms may then serve as ligands for EGFR1.

Epidermal growth factor receptor (EGFR1) plays a pivotal role in the
30 maintenance and repair of epithelial tissue. Following injury in bronchial epithelium, EGFR1 is upregulated in response to ligands acting on it or through

transactivation of the EGFR1 receptor. This results in the increased proliferation of cells and airway remodeling at the point of insult, leading to the repair of the bronchial epithelium (Polosa et. al., 1999, *Am. J. Respir. Cell Mol. Biol.* **20**:914-923; Holgate et. al., 1999, *Clin. Exp. Allergy Suppl* **2**:90-95).

5 In asthma, the bronchial epithelium is highly abnormal, with structural changes involving separation of columnar cells from their basal attachments and functional changes that include increased expression and release of proinflammatory cytokines, growth factors, and mediator-generating enzymes. Beneath this damaged structure are the subepithelial myofibroblasts that have
10 been activated to proliferate. This, in turn, causes excessive matrix deposition leading to abnormal thickening and increased density of the subepithelial basement membrane.

 Immunocytochemical studies have shown that both TGF- β and EGFR1 are highly expressed at the area of injury and that parallel pathways could be
15 operating in the repairing epithelial cells (Puddicombe et. al., 2000, *FASEB J.* **14**:1362-1374). EGFR1 stimulates epithelial repair and TGF- β regulates the production of profibrogenic growth factors and proinflammatory cytokines leading to extracellular matrix synthesis. As EGFR1 is involved in regulating a number of different stages of epithelial repair (survival, migration, proliferation
20 and differentiation), any inhibitory effects that act on the receptor may cause the epithelium to be held in a "state of repair" (Holgate et. al., 1999, *Clin. Exp. Allergy Suppl* **2**:90-95).

 Without wishing to be bound by theory, it is possible that a variant Gene 216 protein induces the epithelium into a continuous "state of repair" by
25 functioning improperly and failing to release its substrate (a member of the neuregulin family) that serves as the ligand for EGFR1. This, in turn, may cause the observed increase in EGFR1 expression. Under these circumstances, the TGF- β pathway remains active, producing a continuous source of proinflammatory products as well as growth factors that drive airway
30 wall remodeling causing bronchial hyperresponsiveness, a phenotype of asthma.

It is also possible that the disintegrin-like domain of Gene 216 plays a role in respiratory diseases. Integrins are a family of heterodimeric transmembrane receptors that mediate cell-cell and cell-extracellular matrix interaction (Hynes, 1992, *Cell* **69**:11). Integrins mediate angiogenesis (Brooks et al., 1994, *Science* **264**:569), which plays a major role in various pathological mechanisms, such as tumor growth, metastasis, diabetic retinopathy, and certain inflammation diseases (Folkman, 1995, *N. Engl. J. Med.* **333**:1757). Disintegrins act as integrin ligands that disrupt cell-matrix interactions (C.P. Blobel and J.M. White, 1992, *Curr. Opin. Cell Biol.* **4**:760-5) and inhibit angiogenesis (C.H. Yeh et al., 1998, *Blood* **92**:3268-3276). Without wishing to be bound by theory, it is possible that the disintegrin-like domain of the Gene 216 polypeptide inhibits angiogenesis in the respiratory system. Gene 216 variants that have partly functional or non-functional disintegrin activity may lack anti-angiogenesis function. These Gene 216 variants may give rise to angiogenesis and inflammation in the respiratory system, a phenotype of asthma.

EXAMPLE 9: Identification of the Mouse Homolog for Gene 216

The mouse ortholog of Gene 216 was identified by TBLASTN analysis of Gene 216 against mouse dbEST. BLAST analysis identified three mouse ESTs that were partially homologous to the human sequence but were not 100% homologous to any known mouse ADAM genes. The three mouse ESTs were 100% identical to a partially sequenced mouse BAC (BAC389B9; Accession Number AF155960). This BAC maps to mouse chromosome 2 in a region that is syntenic to human chromosome 20p13. The 47 kb BAC sequence was analyzed for potential genes using the Genscan gene prediction program (Burge and Karlin, *J. Mol. Biol.*, **268**:78-94). Additional putative exons were identified based on comparison of the human Gene 216 protein to the mouse BAC by TBLASTN. The results identified a mouse gene that contained an ORF of 2124 bp encoding a protein of 707 amino acids. The genomic nucleotide sequence of the mouse homolog is depicted in Figure 21 and the corresponding amino acid sequence is depicted in Figure 22. The mouse

amino acid sequence was analyzed by BLASTP analysis and found to have homology to mouse and human ADAM proteins. The mouse amino acid sequence was aligned against the amino acid sequence of human Gene 216 (BestFit, <http://www.gcg.com>) (Figure 23). The results showed that the mouse and human proteins shared ~70% identity at the amino acid level. This indicated that the mouse sequence was the murine ortholog of human Gene 216.

EXAMPLE 10: Polymorphism Identification

Polymorphisms were identified in the chromosome 20 region and subsequently used in association studies. Most of the data focused on the region of Gene 216.

1. **Single Nucleotide Polymorphism (SNP) Discovery:** An efficient tiered approach was used for mutation analysis. First, PCR assays were developed across exons to include the consensus splice sites. Assays were designed for all exons that contribute to the open reading frame of the gene. This strategy ensured the detection of mutations that would result in the modification of the protein sequence as well as mutations that would be predicted to disrupt mRNA splicing. The identified promoter and putative regulatory element for Gene 216 and a large intronic region were assayed for polymorphisms as well. Second, a total of 77 individuals were tested for polymorphisms using fluorescent SSCP (single strand conformational polymorphism). This sample size provided a 99% power to detect a polymorphism with a frequency of 3% or greater. Briefly, PCR was used to generate templates from asthmatic individuals that showed increased sharing for the 20p13-p12 chromosomal region and contributed towards linkage. Non-asthmatic individuals were used as controls. Enzymatic amplification of Gene 216 was accomplished using PCR with oligonucleotides flanking each exon as well as the putative 5' region. Primers were chosen to amplify each exon as well as 15 or more base pairs within each intron on either side of the splice site. The forward and the reverse primers were labeled with two different dye colors to allow analysis of each strand and confirm variants independently.

Standard PCR assays were utilized for each exon primer pair following optimization. Buffer and cycling conditions were specific to each primer set.

The products were denatured using a formamide dye and electrophoresed on non-denaturing acrylamide gels with varying concentrations of glycerol (at least two different glycerol concentrations).

Primers utilized in fluorescent SSCP experiments to screen coding and non-coding regions of Gene 216 for polymorphisms are provided in Table 8.

Column 1 lists the genes targeted for mutation analysis. Column 2 lists the specific exons analyzed. Column 3 lists the primer names. Columns 4 and 5 list the forward primer sequences and corresponding SEQ ID NOS, respectively. Columns 5 and 6 list the reverse primer sequences and corresponding SEQ ID NOS, respectively

TABLE 8

Gene	Exon	Assay Name	Primer Sequence	SEQ ID NO:	Primer Sequence	SEQ ID NO:
216	216_A	502_216_A_F_503_216_A_R	Ctgcttagaggccgagga	63	agctctgacgagaaccatc	106
216	216_A	1623_216_A_F_1624_216_A_R	Caggagaccacggaagatg	64	ctcgagcgggtggagctg	107
216	216_A	1625_216_A_F_1626_216_A_R	Ttgctgaacctctatcc	65	gagaggaggagagaccgct	108
216	216_B	293_216_B_F_294_216_B_R	Cccctgtgtctcaggtc	66	agtgaacttggtgtcttgg	109
216	216_C	295_216_C_F_296_216_C_R	Gctccacactcttcttgc	67	tgtctctgacccctctctg	110
216	216_D	297_216_D_F_298_216_D_R	Aggcagaggagaagctaat	68	aagaggagggtgtgtgtagg	111
216	216_E	1290_216_E_F_1291_216_E_R	Cctaccacacctccctctt	69	gtgatcaggccactaggggtg	112
216	216_F	299_216_F_F_300_216_F_R	Cctaccctctgacacctta	70	atacgcattccacctcca	113
216	216_G	301_216_G_F_302_216_G_R	aaattctctctggagctgg	71	gaaggcagaatcccggt	114
216	216_H	700_216_H_F_701_216_H_R	cacacctgtgtgaggagaga	72	caccagcacttgcctctc	115
216	216_I	305_216_I_F_306_216_I_R	ccacgaaggaccaccg	73	gggtcngaggccaccac	116
216	216_J	889_216_J_F_890_216_J_R	ctcacgtgggtgctctg	74	gccgtagagcctctgtct	117
216	216_K	891_216_K_F_892_216_K_R	ctctacggccgcagtgac	75	gacgacacaagaacacgag	118
216	216_L	311_216_L_F_312_216_L_R	gtccctccatgcccaatg	76	tgagcggagaggcgaagt	119
216	216_L	313_216_L_F_314_216_L_R	caggtaagtcggctcgc	77	aaacctcacctgaacctt	120
216	216_M	315_216_M_F_316_216_M_R	ctctctgcttcccccac	78	aagggtgctgtgtctct	121
216	216_N	317_216_N_F_318_216_N_R	tctactgtgggaagatggg	79	ccactcagctccactcccta	122
216	216_O	319_216_O_F_320_216_O_R	cccctctacttctcccca	80	ggatitcaaacggcaaggag	123
216	216_P	321_216_P_F_322_216_P_R	gaccttgggttctctatcc	81	gctgagctctgacaggttg	124
216	216_Q	323_216_Q_F_304_216_Q_R	gtgcacctgtcaggactc	82	gaaccgcaggatgagctc	125
216	216_R	325_216_R_F_326_216_R_R	cctgacitctatcacgtgc	83	atatgttcagcaggagacc	126
216	216_S	327_216_S_F_328_216_S_R	ttacctccaccatttctcc	84	gcactcgtgtctctatgata	127
216	216_S	1308_216_S_F_1309_216_S_R	gtggagagggaaggagagag	85	gagcgtttgaatccaggttc	128
216	216_T	1294_216_T_F_1295_216_T_R	cccatgggtgaatttaca	86	cagcaagacaccgcatctac	129
216	216_T	1296_216_T_F_1297_216_T_R	gcagctaggccctacagggtaca	87	gggacagagggaaccattta	130
216	216_T	1298_216_T_F_1299_216_T_R	accacgctatagccaacat	88	ttctctctgttcttccca	131

216	216_T	1300_216_T_F_1301_216_T_R	agggtatagcactgggattgg	89	gtccctgggagcttggtgtgt	132
216	216_T	1302_216_T_F_1303_216_T_R	ccccaggaccactagcttct	90	aggaaccacagagccacacia	133
216	216_T	1304_216_T_F_1305_216_T_R	attgagctggagagtggtcc	91	tgcccttggtgagagtgtagc	134
216	216_T	1306_216_T_F_1307_216_T_R	ttcaagtctctgggtggct	92	ttctggatcactgtgtctc	135
216	216_AA	1619_216_AA_F_1620_216_AA_R	acaaggaccctctaaacgca	93	ttcagcagctgaggaacctct	136
216	216_PQ	1465_216_PQ_F_1466_216_PQ_R	acctctctgtgacaagccag	94	ctgggagctcggtagcaaca	137
216	216_QR	1467_216_QR_F_1468_216_QR_R	gtgtgtctaccgactccag	95	aggccactggaaacctctcl	138
216	216_QR	1469_216_QR_F_1470_216_QR_R	cccaagggtcagagagcag	96	gcagcatggtacagggagctg	139
216	216_QR	1471_216_QR_F_1472_216_QR_R	gcctctctgtccactctct	97	cagctgacagctggtatggga	140
216	216_QR	1473_216_QR_F_1474_216_QR_R	gccaactctctctgcacaaat	98	tgtcagacatggccacagag	141
216	216_QR	1475_216_QR_F_1476_216_QR_R	ttctctgtgacctgggtgt	99	aggctctcttagctgccac	142
216	216_QR	1477_216_QR_F_1478_216_QR_R	atttgggcagagtgatgg	100	agcccttgctcattctgtg	143
216	216_QR	1479_216_QR_F_1480_216_QR_R	ggcagaggagcaaggtgg	101	caaagaaccttggtatgccg	144
216	216_QR	1481_216_QR_F_1482_216_QR_R	atggcttggaatcatcaagg	102	ctcagctccctctctctc	145
216	216_QR	1483_216_QR_F_1484_216_QR_R	tagagagaggaggtgccagc	103	ctgtgtgggacatcttg	146
216	216_RS	1485_216_RS_F_1486_216_RS_R	aaagatggcccaacacagg	104	ggagaaatgggtggaggttaa	147
216	216_ST	1487_216_ST_F_1488_216_ST_R	agaactctcatgagccagc	105	aaagccacagcttctccct	148
216	216_ST	1489_216_ST_F_1490_216_ST_R	aggtttctgggtcagggtta	149	caggatcttggtcatctggac	153
216	216_UP	1463_216_UP_F_1464_216_UP_R	gtagggtggccnagagcagg	150	ctggctgtcacagaagggt	154
216	216_U	1292_216_U_F_1293_216_U_R	tgtggacctagaatggtgagc	151	ctggagcacagtgccaggtta	155
216	216_V	1736_216_V_F_1737_216_V_R	caaagtcacacacaagcgg	152	tttggctgtccctcagttic	156

Once polymorphisms were identified, multiple individuals representative of each SSCP pattern and two genomic controls were sequenced for polymorphism validation and to identify SNPs. The variants detected in the initial set of asthmatic and normal individuals were subject to fluorescent sequencing (ABI) using a standard protocol described by the manufacturer (Perkin Elmer). In cases where SSCP did not identify polymorphisms in Gene 216, sequence information was obtained from 16 individuals that were identical by descent (IBD) in the region, and from 4 controls to ensure that potential polymorphisms were identified.

Primers utilized in DNA sequencing for purposes of confirming polymorphisms detected using fluorescent SSCP are provided in Table 9. Column 1 lists the specific exons sequenced. Column 2 lists the forward primer names, column 3 lists the forward primer sequences, and column 4 lists the corresponding SEQ ID NOS. Column 5 lists the reverse primer names, column 6 lists the reverse primer sequences, and column 7 lists the corresponding SEQ ID NOS.

TABLE 9

Exon	Forward	Forward Seq	SEQ ID NO:	Reverse Name	Reverse Seq	SEQ ID NO:
216 A	MDSeq 101 216 A F	cctctcaggagtagagccc	157	MDSeq 101 216 A R	ccagcagcacttgagcctc	177
216 A	MDSeq 175 216 A F	agcggttctctctctctc	158	MDSeq 175 216 A R	agcctagcctctctctt	178
216 A	MDSeq 213 216 A F	cctctcaggagtagagccc	159	MDSeq 213 216 A R	ccagcagcagcacttga	179
216 A	MDSeq 334 216 A F	atgttaetgagccgaaggt	160	MDSeq 334 216 A R	ccatagacttgagctctc	180
216 B	MDSeq 296 216 B F	cccttccagctctctctt	161	MDSeq 296 216 B R	aaaggttccagcccccata	181
216 C	MDSeq 297 216 C F	caggactgcacacactctga	162	MDSeq 297 216 C R	atcttggtctctccctc	182
216 D	MDSeq 61 216 D F	tccttggttcttccata	163	MDSeq 61 216 D R	ggagagctcttcccca	183
216 E	MDSeq 245 216 E F	agcgagagagagcgtcaat	164	MDSeq 245 216 E R	ggacccagcaggaagctg	184
216 F	MDSeq 57 216 F F	ccctctgcctctcttct	165	MDSeq 57 216 F R	aaccacagctcccgagtg	185
216 G	MDSeq 336 216 G F	ccctgattgccaaggtcttga	166	MDSeq 336 216 G R	ctgctctccctgaagagaa	186
216 H	MDSeq 155 216 H F	agcctcaggtcccaatttt	167	MDSeq 155 216 H R	actgcagcagagcccaag	187
216 I	MDSeq 363 216 I F	agagctctctctctctct	168	MDSeq 363 216 I R	accgaaccttgaacacac	188
216 J	MDSeq 181 216 J F	tgcctcaggtctctcag	169	MDSeq 181 216 J R	tgaagagcagcacaagaac	189
216 K	MDSeq 182 216 K F	tcacgtgtggtctcttga	170	MDSeq 182 216 K R	cnaagctcacacacagccg	190
216 L	MDSeq 106 216 L F	gggttacttccctctcttg	171	MDSeq 106 216 L R	gaacctgagggcaccattt	191
216 M	MDSeq 337 216 M F	cagccltccacccctgg	172	MDSeq 337 216 M R	tggccttagtaattgtgtc	192
216 N	MDSeq 338 216 N F	ggccttctccacccctgg	173	MDSeq 338 216 N R	tggccttagtaattgtgtc	193
216 O	MDSeq 49 216 O F	tccaggatgtgaactctgc	174	MDSeq 49 216 O R	ctgagcagcagctgcaatt	194
216 P	MDSeq 248 216 P F	tgagatgaggtctctctc	175	MDSeq 248 216 P R	agagtgaggtctcaggaaga	195
216 Q	MDSeq 96 216 Q F	ggccttctccacccctgg	176	MDSeq 96 216 Q R	tatctctgagtgtagaggg	196
216 R	MDSeq 50 216 R F	agaggttgatcttgagcaaa	197	MDSeq 50 216 R R	ccaggaacactgtagaggg	219
216 S	MDSeq 262 216 S F	agggcaatacccactcagga	198	MDSeq 262 216 S R	tacctctccagagagcagg	220
216 T	MDSeq 255 216 T F	cccatctgttgaattacata	199	MDSeq 255 216 T R	gcagagagctatgtctctg	221
216 T	MDSeq 256 216 T F	ggcctctggatcctcctac	200	MDSeq 256 216 T R	gcagcagcctttagagatt	222
216 T	MDSeq 257 216 T F	atctcagctgaacatagaggt	201	MDSeq 257 216 T R	ttaactcagcagcagagctc	223
216 T	MDSeq 258 216 T F	ttgttgaactcttctcttg	202	MDSeq 258 216 T R	gacctgattcaagaactctc	224
216 T	MDSeq 358 216 T F	gcagtagcagcattgggaagt	203	MDSeq 358 216 T R	atgttggctatgagcgtct	225
216 T	MDSeq 365 216 T F	actcagctgaacatagaggt	204	MDSeq 365 216 T R	tatctctgagcagcagagtc	226
216 U	MDSeq 244 216 U F	gcagtagcagcattgggaagt	205	MDSeq 244 216 U R	ctgagtgagggagcagagag	227
216 U	MDSeq 292 216 U F	ccaggaaggtgtctatgtct	206	MDSeq 292 216 U R	ctgagtgagggagcagagag	228
216 V	MDSeq 389 216 V F	gggtcattggagggcagag	207	MDSeq 389 216 V R	ccatgagctcgcccaag	229
216 AA	MDSeq 360 216 AA F	gttctctccagattcagat	208	MDSeq 360 216 AA R	atttcaaggctcagtaggg	230
216 PQ	MDSeq 300 216 PQ F	agaatgctctccagagcgtt	209	MDSeq 300 216 PQ R	acttttctcatctctctc	231
216 QR	MDSeq 301 216 QR F	gttttgcaccgactccag	210	MDSeq 301 216 QR R	accaccaaggttcaagagaa	232
216 QR	MDSeq 303 216 QR F	ctgctctcagcctcctc	211	MDSeq 303 216 QR R	tcccaagcagcaggtatgtc	233
216 QR	MDSeq 321 216 QR F	acacagaggttccagtgag	212	MDSeq 321 216 QR R	ctggagatgagagcagc	234
216 QR	MDSeq 322 216 QR F	agcgaggttctgattcaggt	213	MDSeq 322 216 QR R	cttctctctctctctc	235
216 QR	MDSeq 361 216 QR F	ttctcagcgtgaagatgac	214	MDSeq 361 216 QR R	attttgcagaggaagtgag	236
216 QR	MDSeq 362 216 QR F	gccacttctcttcccaaat	215	MDSeq 362 216 QR R	catttctcagcgtctgac	237
216 RS	MDSeq 339 216 RS F	ctgagccagcagacatgatt	216	MDSeq 339 216 RS R	tcagagcctgaggaagatgt	238
216 ST	MDSeq 302 216 ST F	gttagtgagggcaccaggg	217	MDSeq 302 216 ST R	gttctcagctgaggtgag	239
216 UP	MDSeq 359 216 UP F	cctcagtgcccgaggaagtg	218	MDSeq 359 216 UP R	ctggagagctgtagcaaca	240

- Single nucleotide polymorphisms (SNPs) that were identified in Gene
- 5 216 are provided in Table 10. Column 1 lists the SNP numbers (1-48). Column 2 lists the exons that either contain the SNPs or are flanked by intronic sequences that contain the SNPs. Column 3 lists the PMP sites for the SNPs. A "-" denotes polymorphisms which are 5' of the exon that are within the intronic region. The corresponding number is given from the 3' to 5' direction.
- 10 A "+" denotes polymorphisms which are 3' of the exon that are within the intronic region. The number corresponding to the "+" is given from the 5' to 3' direction. Columns 2 and 3, combined, show the SNP names as described herein, e.g., T+1, T+2, etc. Column 4 indicates whether the SNP was detected

- in an exon or intron sequence. Column 5 lists the SNP locations in the Gene 216 genomic sequence of SEQ ID NO:6 (Figure 7). Column 6 lists the SNP reference sequences which illustrate the SNP nucleotide changes with underlining. Column 7 lists the SEQ ID NOs of the SNP reference sequences.
- 5 Column 8 lists the base changes of the SNP sequences. Column 9 lists the amino acid changes resulting from the SNP sequences.

TABLE 10

SNP	Exon	PMP site	Location	Location	Sequence (20nt+ <u>SNP</u> +20nt)	SEQ ID NO:	PMP	AA Change
1	A	-1	intron	4653	GCCTCTGAGACCGACGGGGAGGGACGGCTCGGGCCGGTC	241	A>T	
2	A	-2	intron	4610	CAAGAACCTTCCCAGCGGTTCCTCCTCCTCAGGAGTAG	242	C>A	
3	C	-1	intron	9827	CACCATCTCAGCTCCACACTGTTTCTTGCCCAAGGTCTCGAA	243	C>T	
4	C	-2	intron	9826	CCACCATCTCAGCTCCACACTCTTCTTGCCCAAGGTCTCGA	244	T>A	
5	D	-1	intron	11687	ACAACCTAAGCCATCACCAAGGCTCCTCTCTAGCCCCAAG	245	G>C	
6	D	-2	intron	11661	TGGTGCTTCCCATATTCACATCTCCACAACCTAAGCCATCA	246	T>C	
7	D	1		11912	CAGGATACATAGAAACCCACTACGGCCACAGTGGGCAGCC	247	T>C	Tyr>His
8	F	+1	intron	12545	CCCTCCAAATCAGAAAGAGACAGGAATTACAGGCCCTCGAG	248	A>G	
9	F	1		12411	AGCTGCTCACCTGGAAGGAACTGTGGCCACAGGGATCC	249	A>G	Thr>Ala
10	G	-1	intron	12637	ACTTCCTTCTGGGAGCTGGGGTTGGGGGTACAGGGCTCAAGC	250	G>A	
11	I	1		13197	TTCTGTCAGTGCGCGCGGGGGCTGTGGGCGCAGCGGGCCC	251	G>A	none
12	L	+1	intron	14481	GGTTTCAAGGTGAGGGTTTCGGGGAGCTTGGGAGCCGGCCT	252	G>T	
13	L	-1	intron	14043	CAGAGAAGCGCGGGGGTTGGGGGACTGTCCCTCCATGCCC	253	G>A	
14	L	-2	intron	13988	CCCCCTCTGGGCTCTGCGCGTCTGGCGGCTGTAGCCAAAGC	254	G>A	
15	L	1		14135	CAGCCGCCGCCAGCTGCGCGCTTCTTCCGCAAGGGGGGCG	255	C>T	Ala>Val
16	Q	+1	intron	16158	AGTGGGCTCCCAAGTCAAGCGAGGGGGTGGATCCCTGCCCC	256	A>T	
17	Q	1		15865	TGCTGGCCATGCTCCTCAGCGTCTGCTGCCTCTGTCTCCCA	257	G>A	Val>Ile
18	Q	2		15888	CTGCTGCTCTGCTCCAGGGGCCGGCTGGCTGTGTGTG	258	G>C	
19	QR	+1	intron	16133	GAAGTAGCTTTGAACAGGAGGTTCCAGTGGCCCTCCAGATCA	259	G>T	
20	QR	+3	intron	16361	GCCTCTGTCTACCAAGTTTGGGCCCTTGGCACTTCTCT	260	C>T	
21	QR	+4	intron	16404	ACAAATCACCTCTGTACACCCCTTGAAGTTCCCAATGTGT	261	C>A	
22	QR	+5	intron	16465	TCCATACCACTGGTCAGCTGGGTGCTGGCTGCCCTGTGTG	262	C>T	
23	QR	+6	intron	16486	GGTGCTGGCTGCCCTGTGCGAGGGCCCTGCTTAACCCAG	263	C>T	
24	QR	+7	intron	16936	GGAAATGACAAGGCTTGGGGGATGGGATGGGGACAGTCA	264	G>A	
25	R	+1	intron	17510	AGGGCTCATGCTCCTGCTCTCTCCAGATGGGCAGCACCC	265	C>T	
26	R	+2	intron	17571	GCCCTCCCCAGCCCCAGGGTCTCTGCTGACCATATTAC	266	T>G	
27	R	1		17403	CCTGGGGCGGCTTCAACCCATGGAGTTGGGCCCCACAGCC	267	T>C	Met>Thr
28	R	2		17432	GCCCCACAGCCACTGGACAGCCCTGGCCCTTGGGTGAGTG	268	C>T	Pro>Ser
29	RS	-1	intron	17451	GCCCTGGCCCTGGGTGAGTGAAGGCACCAAGGGGAGGTGG	269	G>T	

30	T	+1	intron	17958	TGCAGCCTGGGGCCCCAGTCC	270	C>A	
31	T	+2	intron	17924	CACTGAGTGAGGATGGGCTC	271	T>C	
32	T	+3	intron	17916	CTGGTCTCACTGAGTGAGG	272	A>G	
33	T	+4	intron	17834	ATGACCTCTTGGTTATCATG	273	G>C	
34	T	1	3' UTR	18833	AGCAAGACACCGCATCTAC	274	G>A	
35	T	2		18787	GGAGGATCACCAGAGGCCAG	275	C>G	
36	T	3	3' UTR	18760	ATCCCAGCACTTGGGAAGCC	276	C>T	
37	T	4	3' UTR	18497	AGCCTGGCTGGCCTCTGCAA	277	A>G	
38	T	5		18476	ACTGAGTCCACACTCCCTTG	278	C>G	
39	T	6		18206	TCCAGGAACCCAGAGCCACT	279	T>C	
40	T	7		18174	TTCTCCCCGAGTGGAGCTTC	280	C>T	
41	T	8		17997	TCCTCACTCTCAGCAGATCA	281	A>T	Gln>His
42	T	-2	intron	19094	CTGAGGACCACACGGGGTGG	282	T>C	
43	T	-4	intron	19160	GGCTGGCAGGCCGAGCCTAG	283	A>G	
44	T	-5	intron	19244	CTTGCTCTGTCACTCCTGCT	284	C>T	
45	U	-1	intron	15423	GTGAGCTCTGCCACCCGACG	285	C>T	
46	V	+1	intron	13859	TGGCGAGGTTACTCTACACG	286	C>T	
47	V	+2	intron	13921	GGCTGCTCACTATTGGGGCC	287	G>T	
48	V	+3	intron	13938	GCCGCATCGTCCCTGTCCC	288	G>A	

Using an in-house program called snp_view; the genomic structure of the gene is diagrammatically shown in Figure 11. The exons are shown to scale and the SNPs are identified by their location along the genomic BAC DNA. The polymorphic sites identified in the Gene 216 genomic sequence are also shown by the underlined nucleotides in Figure 29. The polymorphic sites discovered within the cDNA and the corresponding amino acid position in Gene 216 are underlined in Figure 24. It will be understood by those of skill in the art that the SNPs identified in the Gene 216 genomic sequence can be correlated to the SNP positions identified in the Gene 216 cDNA sequence by aligning the genomic and cDNA sequences.

EXAMPLE 11: Polymorphism Genotyping

Once putative variants were confirmed by sequencing, rapid allele specific assays were designed to type more than 400 individuals (> 200 cases and > 200 controls) for use in the association studies. All coding SNPs (cSNPs) that resulted in an amino acid change were typed. Neutral

polymorphisms were typed if: 1) the polymorphism was present in an exon lacking a cSNP that resulted in an amino acid change; 2) the polymorphism was present in an exon containing a cSNP resulting in an amino acid change but the two polymorphisms were observed to have different frequencies; and
5 3) the polymorphism was in an intronic region adjacent to an exon without a cSNP. If results from the association studies appeared positive, additional neutral polymorphisms were typed. More than 30 allele specific assays from Gene 216 were typed for the case control population (Table 11).

Two types of allele specific assays (ASAs) were used. If the SNP
10 resulted in a mutation that created or abolished a restriction site, restriction fragment length polymorphisms (RFLPs) were obtained from PCR products that spanned the variants, and the RFLPs were analyzed. If the polymorphisms did not result in RFLPs, allele specific oligonucleotide assays were used. For these assays, PCR products that spanned the polymorphism
15 were electrophoresed on agarose gels and transferred to nylon membranes by Southern blotting. Oligomers 16-20 bp in length were designed such that the middle base was specific for each variant. The oligomers were labeled and successively hybridized to the membrane in order to determine genotypes. The specific method used to type each SNP is indicated in Table 11.

20 Table 11 below contains the information relating to the specific assay used. Column 1 lists the SNP designation number. Column 2 lists the specific assay used, either RFLP or ASO. Column 3 lists the enzyme used in the RFLP assay (described below). Columns 4 and 6 list the sequence of the primers used in the ASO assay (described below). Columns 5 and 7 list the
25 corresponding SEQ ID NOS for the primers.

1. RFLP Assay: The amplicon containing the polymorphism was PCR amplified using primers that were used to generate a fragment for sequencing (sequencing primers) or SSCP (SSCP primers). The appropriate population of individuals was PCR amplified in 96 well microtitre plates.

30 Enzymes were purchased from NEB. The restriction cocktail containing the appropriate enzyme for the particular polymorphism is added to the PCR

product. The reaction was incubated at the appropriate temperature according to the manufacturer's recommendations (NEB) for 2-3 hr, followed by a 4°C incubation. After digestion, the reactions were size fractionated using the appropriate agarose gel depending on the assay specifications (2.5%, 3%, or Metaphor, FMC Bioproducts). Gels were electrophoresed in 1 X TBE Buffer at 170 Volts for approximately 2 hr. The gel was illuminated using ultraviolet light and the image was saved as a Kodak 1D file. Using the Kodak 1D image analysis software, the images were scored and the data was exported to Microsoft EXCEL (<http://www.microsoft.com>).

2. ASO assay: The amplicon containing the polymorphism was PCR amplified using primers that were used to generate a fragment for sequencing (sequencing primers) or SSCP (SSCP primers). The appropriate population of individuals was PCR amplified in 96 well microtitre plates and re-arrayed into 384 well microtitre plates using a Tecan Genesis RSP200. The amplified products were loaded onto 2% agarose gels and size fractionated at 150V for 5 min. The DNA was transferred from the gel to Hybond N+ nylon membrane (Amersham-Pharmacia) using a Vacuum blotter (Bio-Rad). The filter containing the blotted PCR products was transferred to a dish containing 300 ml pre-hybridization solution (5 X SSPE (pH 7.4), 2% SDS, 5 X Denhardt's). The filter was incubated in pre-hybridization solution at 40°C for over 1 hr. After pre-hybridization, 10 ml of the pre-hybridization solution and the filter were transferred to a washed glass bottle. The allele specific oligonucleotides (ASO) were designed with the polymorphism in the middle. The size of the oligonucleotide was dependent upon the GC content of the sequence around the polymorphism. Those ASOs that had a G or C polymorphism were designed so that the T_m was between 54-56°C and those that had an A or T variance were designed so that the T_m was between 60-64°C. All oligonucleotides were phosphate free at the 5' end and purchased from GibcoBRL. For each polymorphism, 2 ASOs were designed: one for each variant.

The two ASOs that represented the polymorphism were resuspended

at a concentration of 1 µg/µl and separately end-labeled with γ -ATP³² (6000 Ci/mmol) (NEN) using T4 polynucleotide kinase according to manufacturer recommendations (NEB). The end-labeled products were removed from the unincorporated γ -ATP³² by passing the reactions through Sephadex G-25
5 columns according to manufacturers recommendation (Amersham-Pharmacia). The entire end-labeled product of one ASO was added to the bottle containing the appropriate filter and 10 ml hybridization solution. The hybridization reaction was placed in a rotisserie oven (Hybaid) and left at 40°C for a minimum of 4 hr. The other ASO was stored at -20°C.

10 After the prerequisite hybridization time had elapsed, the filter was removed from the bottle and transferred to 1 L of wash solution (0.1 X SSPE (pH 7.4), 0.1% SDS) pre-warmed to 45°C. After 15 min, the filter was transferred to another L of wash solution (0.1 X SSPE (pH 7.4), 0.1% SDS) pre-warmed to 50°C. After 15 min, the filter was wrapped in Saran, placed in
15 an autoradiograph cassette and an X-ray film (Kodak) placed on top of the filter. Typically, an image would be observed on the film within 1 hr. After an image had been captured on film for the 50°C wash, the process was repeated for wash steps at 55°C, 60°C and 65°C. The image that captured the best result was used.

20 The ASO was removed from the filter by adding 1 L of boiling strip solution (0.1 x SSPE (pH 7.4), 0.1% SDS). This was repeated two more times. After removing the ASO the filter was pre-hybridized in 300 ml pre-hybridization solution (5 X SSPE (pH 7.4), 2% SDS, 5 X Denhardt's) at 40°C for over 1 hr. The second end-labeled ASO corresponding to the other variant
25 was removed from storage at -20°C and thawed at room temperature. The filter was placed into a glass bottle along with 10 ml hybridization solution and the entire end-labeled product of the second ASO. The hybridization reaction was placed in a rotisserie oven (Hybaid, <http://www.hybaid.co.uk>) and left at 40°C for a minimum of 4 hr. After the hybridization, the filter was washed at
30 various temperatures and images captured on film as described above.

The two films that best captured the allele-specific assay with the two

ASOs were converted into digital images by scanning them into Adobe PhotoShop. These images were overlaid against each other in Graphic Converter and then scored.

TABLE 11

SNP	ASA Type	RFLP Enzyme	ASO Primer1	SEQ ID NO:	ASO Primer2	SEQ ID NO:
1	ASO		gcccgcacccccctcg	289	gcccgcctcccctcg	299
2	ASO		ccctctctcttggcgac	290	tcctcctctattggcgacc	300
3	ASO		tcacacactcttcttgc	291	ctccacacttttcttgc	301
4	ASO		gtccacactcttcttgc	292	gtccacactcttcttgc	302
5	ASO		tcaccaagctctctct	293	tcaccaagctctctct	303
6	Alt. Meth					
7	RFLP	XcmI				
8	ASO		cagaagagacaggaatcaca	294	agaagagacgggaatcac	304
9	ASO		tgaagaggaaacctgtggcc	295	tgaagaggagccctgtgg	305
10	ASO					
11	ASO					
12	ASO		gggttccggggagcttg	296	agggttctgtggagcttgg	306
13	ASO		gggttgggggactgtc	297	gggttgggggactgtcc	307
14	ASO		ctctgcgcgtctggcg	298	gtctgcgcactctgggg	308
15	RFLP	BssHII				
16	ASO		agtcaggcgaggagggtgg	309	agtcaggcgaggagggtgg	322
17	ASO		cctcagctcctgtctg	310	ctctcagactcctgtctg	323
18	RFLP	KaI				
19	ASO		aacaggagggttccagtgg	311	gaacaggagttccagtggc	324
20	ASO		accagtitttggcccttt	312	caccagtitttggccctttg	325
21	ASO		ctgtcaccctctgaagt	313	ctgtcaccctgaagtc	326
22	ASO		tcagctgcgggtctgg	314	ggtcagctgtgggtctgg	327
23	RFLP	BstNI				
24	ASO		gccttgggggatgga	315	aggccttgggagatggat	328
25	ASO		tctgcctctccag	316	tcctgcctctccag	329
26	RFLP	BglI				
27	RFLP	NcoI				
28	ASO		actggacagccctggc	317	actggacagctctggc	330
29	ASO					
30	RFLP	Bsu36I				
31	ASO		ctgtgtggcagagccca	318	tgtgtggcaggagccca	331
32	ASO					
33	RFLP	BsaI				
34	Alt. Meth					
35	RFLP	Cac8I				
36	RFLP	MspI				
37	ASO		aattatgtttgttcagaggg	319	attatgtttgttcagaggg	332
38	RFLP	Fnu4III				
39	ASO		gaactctatgtgtgctct	320	ggaactctaatgtgtctctg	333
40	RFLP	TaqI				
41	RFLP	NlaIII				
42	ASO					
43	RFLP	StyI				

44	ASO		ccaagggaggcaggagt	321	ccaaggggaagcaggagta	334
45	RFLP	HinfI				
46	RFLP	BsrI				
47	RFLP	Eco109 I				
48	ASO					

EXAMPLE 12: Association Study Analysis

1. Case-Control Study: In order to determine whether polymorphisms in candidate genes were associated with the asthma phenotype, association studies were performed using a case-control study design. In a well-matched design, the case-control approach is more powerful than the family based transmission disequilibrium test (TDT) (N.E. Morton and A. Collins, 1998, *Proc. Natl. Acad. Sci. USA* **95**:11389-93). Case-control studies are, however, sensitive to population heterogeneity.

To avoid issues of population admixture, which can bias case-control studies, the unaffected controls were collected in both the US and the UK. A total of three hundred controls were collected, 200 in the UK and 100 in the US. Inclusion into the study required that the control individual was negative for asthma, as determined by self-report of never having asthma, had no first degree relatives with asthma, and was negative for eczema and symptoms indicative of atopy within the past 12 months. Data from an abbreviated questionnaire similar to that administered to the affected sib pair families were collected. Results from skin prick tests to 4 common allergens were also collected. The results of the skin prick test were used to select a subset of controls that were most likely to be asthma and atopy negative.

A subset of unrelated cases was selected from the affected sib pair families based on the evidence for linkage at the chromosomal location near a given gene. One affected sib demonstrating identity-by-descent (IBD) at the appropriate marker loci was selected from each family. Since the appropriate cases may vary for each gene in the chromosome 20 region, a larger collection of individuals who were IBD across a larger interval were genotyped, and a subset was used in the analyses. On average, 130 IBD affected individuals and 200 controls were compared for allele and genotype frequencies. This

number provided an 80% power to detect a difference of 5% or greater between the two groups for a rare allele ($\leq 5\%$) at a 0.05 level of significance.

For a common allele (50%), the number provided an 80% power to detect a difference of 10% or more between the two groups.

- 5 For each polymorphism, the frequency of the alleles in the control and case populations was compared using a Fisher exact test. A mutation that increased susceptibility to the disease would be more prevalent in the cases than in the controls, while a protective mutation would be more prevalent in the control group. Similarly, the genotype frequencies of the SNPs were compared
- 10 between cases and controls. P-values for both the allele and genotype were plotted against a coordinate system based on genomic sequence to visualize regions where allelic association was present. A small p-value (or a large value of $-\log(p)$ as plotted in the figures described below) was indicative of an association between the SNPs and the disease phenotype. The analysis was
- 15 repeated for the US and UK population separately to adjust for the possibility of genetic heterogeneity.

2. Association test with individual SNPs: Chromosomal regions harboring asthma susceptibility genes were identified by association studies using the SNP typing data. Two separate phenotypes were used in these
- 20 analyses: asthma and bronchial hyper-responsiveness.

- a. Asthma Phenotype: The significance levels (p-values) for allelic association of all typed SNPs in Gene 216 to the asthma phenotype are plotted in Figure 25 (combined population) and Figure 26 (US and UK populations separately). The most significant result in the combined population
- 25 was observed for Gene 216 exon T+1, where 92.4% of the cases harbored the intronic mutation, while the SNP was present in only 85.2% of the controls ($p = 0.0055$). Six additional SNPs in Gene 216 (T5, QR+7, QR+4, Q2, Q1, and U-1) were significant at the 0.05 level. Frequencies and p-values for SNPs associated with the asthma phenotype in Gene 216 are presented in Tables
- 30 12, 13, and 14 for the combined population and for the UK and US populations, separately.

TABLE 12

Asthma Yes/NO						
Combined US and UK						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	66.5%	215	71.5%	128	0.2029	0.1482
gene216_T_3	8.7%	213	9.5%	131	0.7841	0.6895
gene216_T_4	96.3%	215	98.5%	129	0.1576	0.1513
gene216_T_5	76.7%	217	83.3%	129	0.0420	0.0468
gene216_T_6	77.8%	214	78.4%	125	0.9235	0.9791
gene216_T_7	96.3%	215	98.5%	129	0.1576	0.1513
gene216_T_8	96.5%	211	98.1%	129	0.2528	0.2456
gene216_T_+1	85.2%	216	92.4%	131	0.0055	0.0178
gene216_T_+2	37.3%	209	39.0%	127	0.6825	0.7722
gene216_T_+4	24.4%	215	26.3%	131	0.5886	0.7410
gene216_R_+2	88.3%	217	88.9%	131	0.8076	0.9005
gene216_R_+1	88.7%	191	88.8%	120	1.0000	0.8394
gene216_R_2	9.4%	208	10.8%	125	0.5928	0.7656
gene216_R_1	11.3%	217	11.8%	131	0.9025	0.7483
gene216_QR_+7	78.1%	215	85.7%	129	0.0160	0.0265
gene216_QR_+6	0.5%	216	0.8%	129	0.6323	0.6317
gene216_QR_+5	46.4%	210	48.8%	129	0.5794	0.4165
gene216_QR_+4	51.5%	205	59.9%	126	0.0367	0.1272
gene216_Q_+1	51.2%	206	52.5%	120	0.8075	0.6608
gene216_Q_2	73.7%	217	80.5%	131	0.0432	0.0831
gene216_Q_1	89.5%	209	94.8%	125	0.0213	0.0584
gene216_U_-1	85.0%	217	91.2%	131	0.0184	0.0659
gene216_L_+1	88.7%	213	88.9%	131	1.0000	0.9672
gene216_L_1	99.3%	217	99.6%	131	1.0000	1.0000
gene216_L_-1	88.9%	212	89.2%	130	1.0000	1.0000
gene216_L_-2	92.9%	212	93.1%	131	1.0000	0.9379
gene216_V_+2	71.3%	216	77.1%	129	0.1085	0.2262
gene216_V_+1	96.1%	217	97.2%	125	0.5223	0.5145
gene216_I_1	84.9%	212	85.3%	129	0.9124	1.0000
gene216_G_-1	90.7%	210	91.3%	127	0.8900	0.7683
gene216_F_+1	65.2%	197	70.4%	120	0.1913	0.4109
gene216_F_1	96.8%	217	96.9%	129	1.0000	1.0000
gene216_D_1	0.0%	215	0.4%	131	0.3786	0.3786
gene216_D_-2	0.7%	214	0.8%	127	1.0000	1.0000

TABLE 13

Asthma Yes/No						
UK population						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	65.8%	139	74.3%	101	0.0566	0.1266
gene216_T_3	8.3%	139	9.6%	104	0.6308	0.7329
gene216_T_4	97.1%	140	98.5%	103	0.3689	0.3633
gene216_T_5	75.4%	140	83.3%	102	0.0426	0.0365
gene216_T_6	78.5%	137	80.1%	98	0.7301	0.8875
gene216_T_7	97.5%	138	99.0%	102	0.3129	0.3082
gene216_T_8	97.8%	137	98.5%	102	0.7388	0.7363
gene216_T_+1	86.4%	140	93.8%	104	0.0105	0.0243
gene216_T_+2	37.9%	136	40.5%	100	0.5682	0.8375
gene216_T_+4	25.2%	139	26.0%	104	0.9163	0.6037
gene216_R_+2	87.5%	140	87.5%	104	1.0000	1.0000
gene216_R_+1	86.9%	122	91.1%	95	0.2211	0.4281
gene216_R_2	10.5%	134	8.2%	98	0.4279	0.7007
gene216_R_1	13.2%	140	8.7%	104	0.1473	0.3472
gene216_QR_+7	79.5%	139	86.4%	103	0.0535	0.1362
gene216_QR_+6	0.0%	139	1.0%	103	0.1806	0.1801
gene216_QR_+5	44.4%	133	50.0%	102	0.2273	0.2470
gene216_QR_+4	48.1%	128	59.1%	99	0.0229	0.0730
gene216_Q_+1	53.1%	129	50.5%	97	0.6346	0.5458
gene216_Q_2	72.9%	140	84.6%	104	0.0020	0.0050
gene216_Q_1	89.4%	132	95.1%	101	0.0274	0.0732
gene216_U_-1	86.1%	140	92.3%	104	0.0419	0.0763
gene216_L_+1	87.0%	138	91.8%	104	0.1059	0.2969
gene216_L_1	99.3%	140	99.5%	104	1.0000	1.0000
gene216_L_-1	87.2%	137	92.2%	103	0.0992	0.1655
gene216_L_-2	92.7%	137	92.3%	104	0.8633	1.0000
gene216_V_+2	71.6%	139	79.1%	103	0.0717	0.1519
gene216_V_+1	97.1%	140	98.0%	99	0.7685	0.7655
gene216_I_1	83.7%	138	89.2%	102	0.1094	0.1323
gene216_G_-1	90.2%	137	90.1%	101	1.0000	0.4913
gene216_F_+1	64.1%	128	74.2%	93	0.0295	0.0711
gene216_F_1	97.9%	140	98.0%	102	1.0000	1.0000
gene216_D_1	0.0%	139	0.5%	104	0.4280	0.4280
gene216_D_-2	0.7%	139	1.0%	101	1.0000	1.0000

TABLE 14

Asthma Yes/No						
US population						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	67.8%	76	61.1%	27	0.4053	0.1776
gene216_T_3	9.5%	74	9.3%	27	1.0000	1.0000
gene216_T_4	94.7%	75	98.1%	26	0.4519	0.4404
gene216_T_5	79.2%	77	83.3%	27	0.5583	0.7765
gene216_T_6	76.6%	77	72.2%	27	0.5819	0.6932
gene216_T_7	94.2%	77	96.3%	27	0.7320	0.7241
gene216_T_8	93.9%	74	96.3%	27	0.7308	0.7226
gene216_T_+1	82.9%	76	87.0%	27	0.5262	0.8281
gene216_T_+2	36.3%	73	33.3%	27	0.7416	0.5739
gene216_T_+4	23.0%	76	27.8%	27	0.5795	0.6743
gene216_R_+2	89.6%	77	94.4%	27	0.4127	0.3874
gene216_R_+1	92.0%	69	80.0%	25	0.0334	0.0361
gene216_R_2	7.4%	74	20.4%	27	0.0188	0.0208
gene216_R_1	7.8%	77	24.1%	27	0.0030	0.0055
gene216_QR_+7	75.7%	76	82.7%	26	0.3410	0.0921
gene216_QR_+6	1.3%	77	0.0%	26	1.0000	1.0000
gene216_QR_+5	50.0%	77	44.4%	27	0.5287	0.6337
gene216_QR_+4	57.1%	77	63.0%	27	0.5218	0.4709
gene216_Q_+1	48.1%	77	60.9%	23	0.1345	0.3169
gene216_Q_2	75.3%	77	64.8%	27	0.1571	0.1404
gene216_Q_1	89.6%	77	93.8%	24	0.5726	1.0000
gene216_U_-1	83.1%	77	87.0%	27	0.6654	0.8280
gene216_L_+1	92.0%	75	77.8%	27	0.0116	0.0123
gene216_L_1	99.4%	77	100.0%	27	1.0000	1.0000
gene216_L_-1	92.0%	75	77.8%	27	0.0116	0.0123
gene216_L_-2	93.3%	75	96.3%	27	0.7362	0.5089
gene216_V_+2	70.8%	77	69.2%	26	0.8614	0.8889
gene216_V_+1	94.2%	77	94.2%	26	1.0000	1.0000
gene216_I_1	87.2%	74	70.4%	27	0.0105	0.0074
gene216_G_-1	91.8%	73	96.2%	26	0.3635	0.3440
gene216_F_+1	67.4%	69	57.4%	27	0.2401	0.3270
gene216_F_1	94.8%	77	92.6%	27	0.5136	0.5043
gene216_D_1	0.0%	76	0.0%	27	1.0000	1.0000
gene216_D_-2	0.7%	75	0.0%	26	1.0000	1.0000

- b. Bronchial Hyper-responsiveness: The analyses were repeated using asthmatic children with borderline to severe BHR ($PC_{20} \leq 16$ mg/ml) or $PC_{20}(16)$, as described in the linkage section. First, sibling pairs were identified where both sibs were affected and satisfied this new criteria.
- 5 Of these pairs, one sib was included in the case/control analyses if they showed evidence of linkage at the gene of interest. This phenotype was more restrictive than the Asthma yes/no criteria; hence the number of cases included in the analyses was reduced approximately in half. If the $PC_{20}(16)$ subgroup represented a more genetically homogeneous sample, one expected to see an
- 10 increase in the effect size compared to the one observed in the original set of cases. However, the reduction in sample size could result in estimates that were less accurate and that could obscure a trend in allele frequencies in the control group, the original set of cases and the $PC_{20}(16)$ subgroup. In addition, the reduction in sample size could induce a reduction in power (and increase
- 15 in p values) in spite of the larger effect size.

- The significance levels (p-values) for allelic association of all typed SNPs in Gene 216 to the BHR phenotype are plotted in Figure 27 (combined population) and Figure 28 (US and UK populations separately). Frequencies and p-values for SNPs associated with the BHR phenotype in Gene 216 are
- 20 presented in Tables 15, 16, and 17 for the combined population and for the UK and US populations, separately. Again, multiple SNPs in Gene 216 were associated with the phenotype in each separate population. In the UK population, the most significant SNP was in Gene 216, exon Q2, where 87% of the cases had the mutation compared to 72.9% for the controls ($p = 0.0038$).
- 25 For the US population, the most significant association was found with the SNP in Gene 216 exon R 1, where 28.6% of the cases carried the mutation compared to 7.8% for the controls ($p = 0.0041$).

- In summary, Gene 216 associated with the phenotypes of both asthma and bronchial hyper-responsiveness. Association was found with multiple
- 30 SNPs in both the UK and US populations. The 3' region of the gene, which contains the transmembrane domain, the cytoplasmic domain, and the 3' UTR,

appeared to have the strongest association. Taken together, these data strongly suggested that Gene 216 is an asthma susceptibility gene.

TABLE 15

BHR						
Combined US and UK						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	66.5%	215	67.7%	62	0.8294	0.1358
gene216_T_3	8.7%	213	9.4%	64	0.8592	0.6092
gene216_T_4	96.3%	215	98.4%	62	0.3878	0.3797
gene216_T_5	76.7%	217	79.8%	62	0.5428	0.5315
gene216_T_6	77.8%	214	78.3%	60	1.0000	0.8426
gene216_T_7	96.3%	215	97.7%	64	0.5856	0.5786
gene216_T_8	96.5%	211	97.6%	63	0.7758	0.7721
gene216_T_+1	85.2%	216	90.6%	64	0.1413	0.3117
gene216_T_+2	37.3%	209	41.8%	61	0.3978	0.6939
gene216_T_+4	24.4%	215	26.6%	64	0.6421	0.2498
gene216_R_+2	88.3%	217	88.3%	64	1.0000	0.8975
gene216_R_+1	88.7%	191	89.2%	60	1.0000	0.7540
gene216_R_2	90.6%	208	91.1%	62	1.0000	1.0000
gene216_R_1	11.3%	217	11.7%	64	0.8750	0.7576
gene216_QR_+7	78.1%	215	82.0%	64	0.3876	0.1711
gene216_QR_+6	99.5%	216	100.0%	63	1.0000	1.0000
gene216_QR_+5	46.4%	210	46.8%	63	1.0000	0.5530
gene216_QR_+4	51.5%	205	58.9%	62	0.1521	0.3393
gene216_Q_+1	51.2%	206	51.8%	57	1.0000	0.7632
gene216_Q_2	73.7%	217	79.7%	64	0.2009	0.0664
gene216_Q_1	89.5%	209	94.2%	60	0.1565	0.4299
gene216_U_-1	85.0%	217	89.8%	64	0.1915	0.5304
gene216_L_+1	88.7%	213	89.8%	64	0.8722	0.9410
gene216_L_1	0.7%	217	0.8%	64	1.0000	1.0000
gene216_L_-1	88.9%	212	89.1%	64	1.0000	1.0000
gene216_L_-2	7.1%	212	8.6%	64	0.5661	0.5313
gene216_V_+2	71.3%	216	75.0%	64	0.4343	0.7291
gene216_V_+1	96.1%	217	97.6%	63	0.5874	0.5802
gene216_I_1	84.9%	212	86.7%	64	0.6709	0.8958
gene216_G_-1	9.3%	210	9.5%	63	1.0000	0.9355
gene216_F_+1	65.2%	197	66.7%	57	0.8234	0.3665
gene216_F_1	96.8%	217	97.6%	62	0.7752	0.7715
gene216_D_1	0.0%	215	0.8%	64	0.2294	0.2294
gene216_D_-2	0.7%	214	0.8%	63	1.0000	1.0000

TABLE 16

BHR						
UK population						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	65.8%	139	74.0%	48	0.1635	0.1885
gene216_T_3	8.3%	139	9.0%	50	0.8352	0.6515
gene216_T_4	97.1%	140	98.0%	49	1.0000	1.0000
gene216_T_5	75.4%	140	81.3%	48	0.2641	0.3646
gene216_T_6	78.5%	137	79.4%	46	1.0000	0.9547
gene216_T_7	97.5%	138	98.0%	50	1.0000	1.0000
gene216_T_8	97.8%	137	98.0%	49	1.0000	1.0000
gene216_T_+1	86.4%	140	94.0%	50	0.0454	0.1307
gene216_T_+2	37.9%	136	44.7%	47	0.2715	0.4549
gene216_T_+4	25.2%	139	26.0%	50	0.8938	0.1153
gene216_R_+2	87.5%	140	86.0%	50	0.7290	0.6834
gene216_R_+1	86.9%	122	92.6%	47	0.1838	0.3875
gene216_R_2	89.6%	134	94.8%	48	0.1494	0.4752
gene216_R_1	13.2%	140	7.0%	50	0.1041	0.3226
gene216_QR_+7	79.5%	139	85.0%	50	0.2983	0.3872
gene216_QR_+6	0.0%	139	0.0%	49	1.0000	1.0000
gene216_QR_+5	44.4%	133	49.0%	49	0.4771	0.5020
gene216_QR_+4	48.1%	128	57.3%	48	0.1508	0.2350
gene216_Q_+1	53.1%	129	48.9%	45	0.5407	0.6988
gene216_Q_2	72.9%	140	87.0%	50	0.0038	0.0128
gene216_Q_1	89.4%	132	95.8%	48	0.0613	0.1924
gene216_U_-1	86.1%	140	93.0%	50	0.0752	0.2087
gene216_L_+1	87.0%	138	94.0%	50	0.0638	0.2367
gene216_L_1	0.7%	140	1.0%	50	1.0000	1.0000
gene216_L_-1	87.2%	137	93.0%	50	0.1400	0.3796
gene216_L_-2	7.3%	137	9.0%	50	0.6623	0.5686
gene216_V_+2	71.6%	139	79.0%	50	0.1860	0.3615
gene216_V_+1	97.1%	140	98.0%	49	1.0000	1.0000
gene216_I_1	83.7%	138	91.0%	50	0.0952	0.2406
gene216_G_-1	9.9%	137	10.2%	49	1.0000	0.9269
gene216_F_+1	64.1%	128	73.3%	43	0.1466	0.2885
gene216_F_1	97.9%	140	97.9%	48	1.0000	1.0000
gene216_D_1	0.0%	139	1.0%	50	0.2646	0.2646
gene216_D_-2	0.7%	139	1.0%	49	1.0000	1.0000

TABLE 17

BHR						
US population						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	67.8%	76	46.4%	14	0.0514	0.0409
gene216_T_3	9.5%	74	10.7%	14	0.7369	1.0000
gene216_T_4	94.7%	75	100.0%	13	0.6065	0.5986
gene216_T_5	79.2%	77	75.0%	14	0.6206	0.6767
gene216_T_6	76.6%	77	75.0%	14	0.8130	0.7738
gene216_T_7	94.2%	77	96.4%	14	1.0000	1.0000
gene216_T_8	93.9%	74	96.4%	14	1.0000	1.0000
gene216_T_+1	82.9%	76	78.6%	14	0.5937	0.6635
gene216_T_+2	36.3%	73	32.1%	14	0.8300	1.0000
gene216_T_+4	23.0%	76	28.6%	14	0.6296	0.7242
gene216_R_+2	89.6%	77	96.4%	14	0.4778	0.4545
gene216_R_+1	92.0%	69	76.9%	13	0.0321	0.0452
gene216_R_2	92.6%	74	78.6%	14	0.0333	0.0469
gene216_R_1	7.8%	77	28.6%	14	0.0041	0.0072
gene216_QR_+7	75.7%	76	71.4%	14	0.6391	0.2476
gene216_QR_+6	98.7%	77	100.0%	14	1.0000	1.0000
gene216_QR_+5	50.0%	77	39.3%	14	0.3130	0.4007
gene216_QR_+4	57.1%	77	64.3%	14	0.5371	0.8691
gene216_Q_+1	48.1%	77	62.5%	12	0.2724	0.4060
gene216_Q_2	75.3%	77	53.6%	14	0.0233	0.0331
gene216_Q_1	89.6%	77	87.5%	12	0.7250	0.5718
gene216_U_-1	83.1%	77	78.6%	14	0.5910	0.6593
gene216_L_+1	92.0%	75	75.0%	14	0.0149	0.0227
gene216_L_1	0.6%	77	0.0%	14	1.0000	1.0000
gene216_L_-1	92.0%	75	75.0%	14	0.0149	0.0227
gene216_L_-2	6.7%	75	7.1%	14	1.0000	1.0000
gene216_V_+2	70.8%	77	60.7%	14	0.3730	0.2711
gene216_V_+1	94.2%	77	96.4%	14	1.0000	1.0000
gene216_I_1	87.2%	74	71.4%	14	0.0455	0.0463
gene216_G_-1	8.2%	73	7.1%	14	1.0000	1.0000
gene216_F_+1	67.4%	69	46.4%	14	0.0510	0.0865
gene216_F_1	94.8%	77	96.4%	14	1.0000	1.0000
gene216_D_1	0.0%	76	0.0%	14	1.0000	1.0000
gene216_D_-2	0.7%	75	0.0%	14	1.0000	1.0000

EXAMPLE 13: Haplotype analyses

In addition to the analysis of individual SNPs, haplotype frequencies between the case and control groups were also compared. The haplotypes were constructed using a maximum likelihood approach. Since existing
 5 software for predicting haplotypes is unable to utilize individuals with missing data, a program was developed to make use of all individuals and, hence, provide more accurate haplotype frequency estimates. Haplotype analysis based on multiple SNPs in a gene is expected to provide increased evidence for an association between a given phenotype and that gene if all haplotyped
 10 SNPs are involved in the characterization of the phenotype. In other words, allelic variation involving those haplotyped SNPs are expected to be associated with different risks or susceptibilities toward the phenotype.

1. Asthma phenotype: The estimated frequency of each haplotype was compared between cases and controls by a permutation test. An overall
 15 comparison of the distribution of all haplotypes between the two groups was also performed. In Tables 18, 19 and 20 the haplotype analysis (2-at-a-time) for all SNPs in Gene 216 is presented for the combined, the UK and the US populations, respectively. The diagonal entries represent the single SNP p-values, while the other entries are the p-values for a test of association
 20 between the asthma phenotype and the haplotypes defined by the 2 SNPs listed on the horizontal and vertical axes. The frequency of the individual SNPs in the cases and controls are shown at the bottom of the tables. Colored cells indicate p-values that were statistically significant (light gray: 0.01 to 0.05, dark gray: 0.001 to 0.0099, black: < 0.001). As seen in Table 18, haplotypes
 25 defined by SNPs T5 & T8, SNPs T+2 & QR+4, T5 & T7 and SNPs T4 & T5, yielded highly significant p-values of 0.00039, 0.000042, 0.00056 and 0.00042 respectively, which were more significant than the analysis of these SNPs alone (T4 p = 0.16; T5 p = 0.04; T7 p = 0.16; T8 p = 0.25; T+2 p = 0.68; QR+4 p = 0.04). These associations were also more significant than the one
 30 observed for the single SNP T+1 reported above. In the UK population, the most significant association was found in Gene 216 (Table 19) with five

- haplotypes significant at the 0.001 level (SNPs T+2 & QR+4, $p = 0.000021$; QR+5 & QR+4, $p = 0.00051$; QR+4 & Q+1 $p = 0.00066$; QR+6 & Q2, $p = 0.00062$; and QR+4 & Q2, $p = 0.00023$). Forty four haplotypes were significant at the 0.01 level in Gene 216 (Table 19) in the UK population. In the US
- 5 population, numerous haplotypes were significant at the 0.01 level for Genes 216 (Table 20).

TABLE 18

	216 I 2	216 I 3	216 I 4	216 I 5	216 I 6	216 I 7	216 I 8	216 I 9	216 I 10	216 I 11	216 I 12	216 I 13	216 I 14	216 I 15	216 I 16	216 I 17	216 I 18	216 I 19	216 I 20	216 I 21	216 I 22	216 I 23	216 I 24	216 I 25	216 I 26	216 I 27	216 I 28	216 I 29	216 I 30	216 I 31	216 I 32	216 I 33	216 I 34	216 I 35	216 I 36	216 I 37	216 I 38	216 I 39	216 I 40	216 I 41	216 I 42	216 I 43	216 I 44	216 I 45	216 I 46	216 I 47	216 I 48	216 I 49	216 I 50	216 I 51	216 I 52	216 I 53	216 I 54	216 I 55	216 I 56	216 I 57	216 I 58	216 I 59	216 I 60	216 I 61	216 I 62	216 I 63	216 I 64	216 I 65	216 I 66	216 I 67	216 I 68	216 I 69	216 I 70	216 I 71	216 I 72	216 I 73	216 I 74	216 I 75	216 I 76	216 I 77	216 I 78	216 I 79	216 I 80	216 I 81	216 I 82	216 I 83	216 I 84	216 I 85	216 I 86	216 I 87	216 I 88	216 I 89	216 I 90	216 I 91	216 I 92	216 I 93	216 I 94	216 I 95	216 I 96	216 I 97	216 I 98	216 I 99	216 I 100	216 I 101	216 I 102	216 I 103	216 I 104	216 I 105	216 I 106	216 I 107	216 I 108	216 I 109	216 I 110	216 I 111	216 I 112	216 I 113	216 I 114	216 I 115	216 I 116	216 I 117	216 I 118	216 I 119	216 I 120	216 I 121	216 I 122	216 I 123	216 I 124	216 I 125	216 I 126	216 I 127	216 I 128	216 I 129	216 I 130	216 I 131	216 I 132	216 I 133	216 I 134	216 I 135	216 I 136	216 I 137	216 I 138	216 I 139	216 I 140	216 I 141	216 I 142	216 I 143	216 I 144	216 I 145	216 I 146	216 I 147	216 I 148	216 I 149	216 I 150	216 I 151	216 I 152	216 I 153	216 I 154	216 I 155	216 I 156	216 I 157	216 I 158	216 I 159	216 I 160	216 I 161	216 I 162	216 I 163	216 I 164	216 I 165	216 I 166	216 I 167	216 I 168	216 I 169	216 I 170	216 I 171	216 I 172	216 I 173	216 I 174	216 I 175	216 I 176	216 I 177	216 I 178	216 I 179	216 I 180	216 I 181	216 I 182	216 I 183	216 I 184	216 I 185	216 I 186	216 I 187	216 I 188	216 I 189	216 I 190	216 I 191	216 I 192	216 I 193	216 I 194	216 I 195	216 I 196	216 I 197	216 I 198	216 I 199	216 I 200	216 I 201	216 I 202	216 I 203	216 I 204	216 I 205	216 I 206	216 I 207	216 I 208	216 I 209	216 I 210	216 I 211	216 I 212	216 I 213	216 I 214	216 I 215	216 I 216	216 I 217	216 I 218	216 I 219	216 I 220	216 I 221	216 I 222	216 I 223	216 I 224	216 I 225	216 I 226	216 I 227	216 I 228	216 I 229	216 I 230	216 I 231	216 I 232	216 I 233	216 I 234	216 I 235	216 I 236	216 I 237	216 I 238	216 I 239	216 I 240	216 I 241	216 I 242	216 I 243	216 I 244	216 I 245	216 I 246	216 I 247	216 I 248	216 I 249	216 I 250	216 I 251	216 I 252	216 I 253	216 I 254	216 I 255	216 I 256	216 I 257	216 I 258	216 I 259	216 I 260	216 I 261	216 I 262	216 I 263	216 I 264	216 I 265	216 I 266	216 I 267	216 I 268	216 I 269	216 I 270	216 I 271	216 I 272	216 I 273	216 I 274	216 I 275	216 I 276	216 I 277	216 I 278	216 I 279	216 I 280	216 I 281	216 I 282	216 I 283	216 I 284	216 I 285	216 I 286	216 I 287	216 I 288	216 I 289	216 I 290	216 I 291	216 I 292	216 I 293	216 I 294	216 I 295	216 I 296	216 I 297	216 I 298	216 I 299	216 I 300	216 I 301	216 I 302	216 I 303	216 I 304	216 I 305	216 I 306	216 I 307	216 I 308	216 I 309	216 I 310	216 I 311	216 I 312	216 I 313	216 I 314	216 I 315	216 I 316	216 I 317	216 I 318	216 I 319	216 I 320	216 I 321	216 I 322	216 I 323	216 I 324	216 I 325	216 I 326	216 I 327	216 I 328	216 I 329	216 I 330	216 I 331	216 I 332	216 I 333	216 I 334	216 I 335	216 I 336	216 I 337	216 I 338	216 I 339	216 I 340	216 I 341	216 I 342	216 I 343	216 I 344	216 I 345	216 I 346	216 I 347	216 I 348	216 I 349	216 I 350	216 I 351	216 I 352	216 I 353	216 I 354	216 I 355	216 I 356	216 I 357	216 I 358	216 I 359	216 I 360	216 I 361	216 I 362	216 I 363	216 I 364	216 I 365	216 I 366	216 I 367	216 I 368	216 I 369	216 I 370	216 I 371	216 I 372	216 I 373	216 I 374	216 I 375	216 I 376	216 I 377	216 I 378	216 I 379	216 I 380	216 I 381	216 I 382	216 I 383	216 I 384	216 I 385	216 I 386	216 I 387	216 I 388	216 I 389	216 I 390	216 I 391	216 I 392	216 I 393	216 I 394	216 I 395	216 I 396	216 I 397	216 I 398	216 I 399	216 I 400	216 I 401	216 I 402	216 I 403	216 I 404	216 I 405	216 I 406	216 I 407	216 I 408	216 I 409	216 I 410	216 I 411	216 I 412	216 I 413	216 I 414	216 I 415	216 I 416	216 I 417	216 I 418	216 I 419	216 I 420	216 I 421	216 I 422	216 I 423	216 I 424	216 I 425	216 I 426	216 I 427	216 I 428	216 I 429	216 I 430	216 I 431	216 I 432	216 I 433	216 I 434	216 I 435	216 I 436	216 I 437	216 I 438	216 I 439	216 I 440	216 I 441	216 I 442	216 I 443	216 I 444	216 I 445	216 I 446	216 I 447	216 I 448	216 I 449	216 I 450	216 I 451	216 I 452	216 I 453	216 I 454	216 I 455	216 I 456	216 I 457	216 I 458	216 I 459	216 I 460	216 I 461	216 I 462	216 I 463	216 I 464	216 I 465	216 I 466	216 I 467	216 I 468	216 I 469	216 I 470	216 I 471	216 I 472	216 I 473	216 I 474	216 I 475	216 I 476	216 I 477	216 I 478	216 I 479	216 I 480	216 I 481	216 I 482	216 I 483	216 I 484	216 I 485	216 I 486	216 I 487	216 I 488	216 I 489	216 I 490	216 I 491	216 I 492	216 I 493	216 I 494	216 I 495	216 I 496	216 I 497	216 I 498	216 I 499	216 I 500	216 I 501	216 I 502	216 I 503	216 I 504	216 I 505	216 I 506	216 I 507	216 I 508	216 I 509	216 I 510	216 I 511	216 I 512	216 I 513	216 I 514	216 I 515	216 I 516	216 I 517	216 I 518	216 I 519	216 I 520	216 I 521	216 I 522	216 I 523	216 I 524	216 I 525	216 I 526	216 I 527	216 I 528	216 I 529	216 I 530	216 I 531	216 I 532	216 I 533	216 I 534	216 I 535	216 I 536	216 I 537	216 I 538	216 I 539	216 I 540	216 I 541	216 I 542	216 I 543	216 I 544	216 I 545	216 I 546	216 I 547	216 I 548	216 I 549	216 I 550	216 I 551	216 I 552	216 I 553	216 I 554	216 I 555	216 I 556	216 I 557	216 I 558	216 I 559	216 I 560	216 I 561	216 I 562	216 I 563	216 I 564	216 I 565	216 I 566	216 I 567	216 I 568	216 I 569	216 I 570	216 I 571	216 I 572	216 I 573	216 I 574	216 I 575	216 I 576	216 I 577	216 I 578	216 I 579	216 I 580	216 I 581	216 I 582	216 I 583	216 I 584	216 I 585	216 I 586	216 I 587	216 I 588	216 I 589	216 I 590	216 I 591	216 I 592	216 I 593	216 I 594	216 I 595	216 I 596	216 I 597	216 I 598	216 I 599	216 I 600	216 I 601	216 I 602	216 I 603	216 I 604	216 I 605	216 I 606	216 I 607	216 I 608	216 I 609	216 I 610	216 I 611	216 I 612	216 I 613	216 I 614	216 I 615	216 I 616	216 I 617	216 I 618	216 I 619	216 I 620	216 I 621	216 I 622	216 I 623	216 I 624	216 I 625	216 I 626	216 I 627	216 I 628	216 I 629	216 I 630	216 I 631	216 I 632	216 I 633	216 I 634	216 I 635	216 I 636	216 I 637	216 I 638	216 I 639	216 I 640	216 I 641	216 I 642	216 I 643	216 I 644	216 I 645	216 I 646	216 I 647	216 I 648	216 I 649	216 I 650	216 I 651	216 I 652	216 I 653	216 I 654	216 I 655	216 I 656	216 I 657	216 I 658	216 I 659	216 I 660	216 I 661	216 I 662	216 I 663	216 I 664	216 I 665	216 I 666	216 I 667	216 I 668	216 I 669	216 I 670	216 I 671	216 I 672	216 I 673	216 I 674	216 I 675	216 I 676	216 I 677	216 I 678	216 I 679	216 I 680	216 I 681	216 I 682	216 I 683	216 I 684	216 I 685	216 I 686	216 I 687	216 I 688	216 I 689	216 I 690	216 I 691	216 I 692	216 I 693	216 I 694	216 I 695	216 I 696	216 I 697	216 I 698	216 I 699	216 I 700	216 I 701	216 I 702	216 I 703	216 I 704	216 I 705	216 I 706	216 I 707	216 I 708	216 I 709	216 I 710	216 I 711	216 I 712	216 I 713	216 I 714	216 I 715	216 I 716	216 I 717	216 I 718	216 I 719	216 I 720	216 I 721	216 I 722	216 I 723	216 I 724	216 I 725	216 I 726	216 I 727	216 I 728	216 I 729	216 I 730	216 I 731	216 I 732	216 I 733	216 I 734	216 I 735	216 I 736	216 I 737	216 I 738	216 I 739	216 I 740	216 I 741	216 I 742	216 I 743	216 I 744	216 I 745	216 I 746	216 I 747	216 I 748	216 I 749	216 I 750	216 I 751	216 I 752	216 I 753	216 I 754	216 I 755	216 I 756	216 I 757	216 I 758	216 I 759	216 I 760	216 I 761	216 I 762	216 I 763	216 I 764	216 I 765	216 I 766	216 I 767	216 I 768	216 I 769	216 I 770	216 I 771	216 I 772	216 I 773	216 I 774	216 I 775	216 I 776	216 I 777	216 I 778	216 I 779	216 I 780	216 I 781	216 I 782	216 I 783	216 I 784	216 I 785	216 I 786	216 I 787	216 I 788	216 I 789	216 I 790	216 I 791	216 I 792	216 I 793	216 I 794	216 I 795	216 I 796	216 I 797	216 I 798	216 I 799	216 I 800	216 I 801	216 I 802	216 I 803	216 I 804	216 I 805	216 I 806	216 I 807	216 I 808	216 I 809	216 I 810	216 I 811	216 I 812	216 I 813	216 I 814	216 I 815	216 I 816	216 I 817	216 I 818	216 I 819	216 I 820	216 I 821	216 I 822	216 I 823	216 I 824	216 I 825	216 I 826	216 I 827	216 I 828	216 I 829	216 I 830	216 I 831	216 I 832	216 I 833	216 I 834	216 I 835	216 I 836	216 I 837	216 I 838	216 I 839	216 I 840	216 I 841	216 I 842	216 I 843	216 I 844	216 I 845	216 I 846	216 I 847	216 I 848	216 I 849	216 I 850	216 I 851	216 I 852	216 I 853	216 I 854	216 I 855	216 I 856	216 I 857	216 I 858	216 I 859	216 I 860	216 I 861	216 I 862	216 I 863	216 I 864	216 I 865	216 I 866	216 I 867	216 I 868	216 I 869	216 I 870	216 I 871	216 I 872	216 I 873	216 I 874	216 I 875	216 I 876	216 I 877	216 I 878	216 I 879	216 I 880	216 I 881	216 I 882	216 I 883	216 I 884	216 I 885	216 I 886	216 I 887	216 I 888	216 I 889	216 I 890	216 I 891	216 I 892	216 I 893	216 I 894	216 I 895	216 I 896	216 I 897	216 I 898	216 I 899	216 I 900	216 I 901	216 I 902	216 I 903	216 I 904	216 I 905	216 I 906	216 I 907	216 I 908	216 I 909	216 I 910	216 I 911	216 I 912	216 I 913	216 I 914	216 I 915	216 I 916	216 I 917	216 I 918	216 I 919	216 I 920	216 I 921	216 I 922	216 I 923	216 I 924	216 I 925	216 I 926	216 I 927	216 I 928	216 I 929	216 I 930	216 I 931	216 I 932	216 I 933	216 I 934	216 I 935	216 I 936	216 I 937	216 I 938	216 I 939	216 I 940	216 I 941	216 I 942	216 I 943	216 I 944	216 I 945	216 I 946	216 I 947	216 I 948	216 I 949	216 I 950	216 I 951	216 I 952	216 I 953	216 I 954	216 I 955	216 I 956	216 I 957	216 I 958	216 I 959	216 I 960	216 I 961	
--	---------	---------	---------	---------	---------	---------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	--

TABLE 18 (CONT'D)

	216.Q.-1	216.Q.2	216.Q.1	216.U.-1	216.L.-1	216.L.-2	216.V.-2	216.V.-1	216.L.1	216.G.-1	216.F.-1	216.F.-1	216.D.-1	216.D.-2
216.L.-2	0.7775	0.2375	0.1515	0.0818	0.4473	0.3851	0.2018	0.4382	0.4252	0.5576	0.4529	0.5379	0.1317	0.3817 216.L.2
216.L.-3	0.8932	0.1511	0.0315	0.0574	0.2518	0.5855	0.5855	0.2331	0.3133	0.9839	0.281	0.4089	0.4333	0.3258 216.L.3
216.L.-4	0.307	0.092	0.0147	0.063	0.2503	0.2765	0.3028	0.1003	0.4414	0.3837	0.2559	0.1684	0.064	0.2689 216.L.4
216.L.-5	0.1626	0.0816	0.1025	0.0361	0.1693	0.1506	0.1146	0.0899	0.0708	0.1449	0.1469	0.0164	0.0292	0.1550 216.L.5
216.L.-6	0.16167	0.1862	0.044	0.1048	0.9885	0.2738	0.3852	0.7482	0.9555	0.9741	0.5124	0.6764	0.532	0.3411 216.L.6
216.L.-7	0.5334	0.0571	0.0184	0.0424	0.2339	0.2698	0.2587	0.1161	0.3118	0.3385	0.1307	0.0182	0.0594	0.2707 216.L.7
216.L.-8	0.4772	0.0951	0.0305	0.051	0.4859	0.4205	0.4468	0.4638	0.2146	0.4287	0.2304	0.0388	0.1513	0.3918 216.L.8
216.L.-1	0.0286	0.0186	0.0168	0.0062	0.0177	0.0137	0.0161	0.033	0.0347	0.0346	0.0254	0.0254	0.0254	0.0228 216.L.1
216.L.-2	0.9755	0.243	0.0533	0.0617	0.3586	0.8847	0.1193	0.0527	0.7459	0.9803	0.3267	0.9157	0.4278	0.8518 216.L.2
216.L.-3	0.8364	0.1748	0.0573	0.0324	0.8431	0.4011	0.5552	0.3519	0.803	0.9044	0.2859	0.5885	0.3754	0.8453 216.L.3
216.L.-4	0.8907	0.0886	0.0278	0.0039	0.5954	0.2592	0.3259	0.1872	0.705	0.9578	0.5821	0.9853	0.4631	0.8772 216.L.4
216.L.-5	0.8471	0.1123	0.0448	0.0497	0.8427	0.5391	0.3593	0.3227	0.3138	0.9533	0.426	0.9396	0.5725	0.9732 216.L.5
216.L.-6	0.8914	0.061	0.0599	0.0546	0.9526	0.8944	0.7497	0.8555	0.461	0.9581	0.3036	0.9387	0.3937	0.831 216.L.6
216.L.-7	0.8095	0.0549	0.0512	0.0598	0.6993	0.844	0.5973	0.8528	0.2512	0.9884	0.372	0.9825	0.5493	0.9466 216.L.7
216.L.-8	0.0096	0.045	0.0421	0.0506	0.0508	0.0531	0.0448	0.0146	0.0920	0.1113	0.0792	0.023	0.0123	0.0568 216.L.8
216.QR.-7	0.8136	0.2301	0.0896	0.0915	0.9711	0.7035	0.3535	0.7824	0.5544	0.8363	0.4564	0.3706	0.5073	0.8373 216.QR.-7
216.QR.-6	0.418	0.0287	0.0255	0.0289	0.8441	0.6629	0.3208	0.3533	0.316	0.8396	0.1319	0.1467	0.3787	0.8623 216.QR.-6
216.QR.-5	0.0343	0.0265	0.0335	0.0345	0.0739	0.0418	0.0562	0.0378	0.1122	0.1189	0.0580	0.1038	0.0368	0.1518 216.QR.-5
216.Q.-1	0.8075	0.0188	0.0299	0.0391	0.8294	0.5062	0.3533	0.316	0.2531	0.485	0.0801	0.1038	0.0368	0.1518 216.Q.-1
216.Q.2	0.8432	0.0577	0.052	0.0213	0.0408	0.1089	0.042	0.0333	0.0765	0.2605	0.0546	0.0546	0.0313	0.1983 216.Q.2
216.Q.1	0.0517	0.0517	0.0517	0.0517	0.0537	0.0533	0.0519	0.068	0.0344	0.0476	0.0589	0.0554	0.0135	0.0559 216.Q.1
216.L.-1	0.0184	0.0184	0.0184	0.0184	0.0539	0.0554	0.0394	0.0778	0.1373	0.0565	0.049	0.0908	0.0728	0.0559 216.L.-1
216.L.-2	0.1548	0.1548	0.1548	0.1548	0.548	0.8826	0.2389	0.7837	0.7182	0.82	0.349	0.9301	0.4305	0.7679 216.L.-2
216.L.-3	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.L.-3
216.L.-4	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.L.-4
216.L.-5	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.L.-5
216.L.-6	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.L.-6
216.L.-7	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.L.-7
216.L.-8	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.L.-8
216.G.-1	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.G.-1
216.F.-1	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.F.-1
216.D.-1	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.D.-1
216.D.-2	0.0929	0.0929	0.0929	0.0929	0.3991	0.3991	0.2175	0.3505	0.3502	0.3502	0.3559	0.387	0.5769	0.4721 216.D.-2
CNTL	51.20%	73.70%	89.50%	85.00%	86.70%	99.30%	98.90%	98.10%	98.10%	98.10%	98.10%	98.10%	98.10%	98.10%
CASE	82.50%	80.50%	84.80%	91.20%	88.90%	99.60%	99.20%	97.10%	97.20%	97.20%	97.20%	97.20%	97.20%	97.20%

TABLE 19

218_I,2	218_I,3	218_I,4	218_I,5	218_I,6	218_I,7	218_I,8	218_I,9	218_I,10	218_I,11	218_I,12	218_I,13	218_I,14	218_I,15	218_I,16	218_I,17	218_I,18	218_I,19	218_I,20	218_I,21	218_I,22	218_I,23	218_I,24	218_I,25	218_I,26	218_I,27	218_I,28	218_I,29	218_I,30	218_I,31	218_I,32	218_I,33	218_I,34	218_I,35	218_I,36	218_I,37	218_I,38	218_I,39	218_I,40	218_I,41	218_I,42	218_I,43	218_I,44	218_I,45	218_I,46	218_I,47	218_I,48	218_I,49	218_I,50	218_I,51	218_I,52	218_I,53	218_I,54	218_I,55	218_I,56	218_I,57	218_I,58	218_I,59	218_I,60	218_I,61	218_I,62	218_I,63	218_I,64	218_I,65	218_I,66	218_I,67	218_I,68	218_I,69	218_I,70	218_I,71	218_I,72	218_I,73	218_I,74	218_I,75	218_I,76	218_I,77	218_I,78	218_I,79	218_I,80	218_I,81	218_I,82	218_I,83	218_I,84	218_I,85	218_I,86	218_I,87	218_I,88	218_I,89	218_I,90	218_I,91	218_I,92	218_I,93	218_I,94	218_I,95	218_I,96	218_I,97	218_I,98	218_I,99	218_I,100	218_I,101	218_I,102	218_I,103	218_I,104	218_I,105	218_I,106	218_I,107	218_I,108	218_I,109	218_I,110	218_I,111	218_I,112	218_I,113	218_I,114	218_I,115	218_I,116	218_I,117	218_I,118	218_I,119	218_I,120	218_I,121	218_I,122	218_I,123	218_I,124	218_I,125	218_I,126	218_I,127	218_I,128	218_I,129	218_I,130	218_I,131	218_I,132	218_I,133	218_I,134	218_I,135	218_I,136	218_I,137	218_I,138	218_I,139	218_I,140	218_I,141	218_I,142	218_I,143	218_I,144	218_I,145	218_I,146	218_I,147	218_I,148	218_I,149	218_I,150	218_I,151	218_I,152	218_I,153	218_I,154	218_I,155	218_I,156	218_I,157	218_I,158	218_I,159	218_I,160	218_I,161	218_I,162	218_I,163	218_I,164	218_I,165	218_I,166	218_I,167	218_I,168	218_I,169	218_I,170	218_I,171	218_I,172	218_I,173	218_I,174	218_I,175	218_I,176	218_I,177	218_I,178	218_I,179	218_I,180	218_I,181	218_I,182	218_I,183	218_I,184	218_I,185	218_I,186	218_I,187	218_I,188	218_I,189	218_I,190	218_I,191	218_I,192	218_I,193	218_I,194	218_I,195	218_I,196	218_I,197	218_I,198	218_I,199	218_I,200	218_I,201	218_I,202	218_I,203	218_I,204	218_I,205	218_I,206	218_I,207	218_I,208	218_I,209	218_I,210	218_I,211	218_I,212	218_I,213	218_I,214	218_I,215	218_I,216	218_I,217	218_I,218	218_I,219	218_I,220	218_I,221	218_I,222	218_I,223	218_I,224	218_I,225	218_I,226	218_I,227	218_I,228	218_I,229	218_I,230	218_I,231	218_I,232	218_I,233	218_I,234	218_I,235	218_I,236	218_I,237	218_I,238	218_I,239	218_I,240	218_I,241	218_I,242	218_I,243	218_I,244	218_I,245	218_I,246	218_I,247	218_I,248	218_I,249	218_I,250	218_I,251	218_I,252	218_I,253	218_I,254	218_I,255	218_I,256	218_I,257	218_I,258	218_I,259	218_I,260	218_I,261	218_I,262	218_I,263	218_I,264	218_I,265	218_I,266	218_I,267	218_I,268	218_I,269	218_I,270	218_I,271	218_I,272	218_I,273	218_I,274	218_I,275	218_I,276	218_I,277	218_I,278	218_I,279	218_I,280	218_I,281	218_I,282	218_I,283	218_I,284	218_I,285	218_I,286	218_I,287	218_I,288	218_I,289	218_I,290	218_I,291	218_I,292	218_I,293	218_I,294	218_I,295	218_I,296	218_I,297	218_I,298	218_I,299	218_I,300	218_I,301	218_I,302	218_I,303	218_I,304	218_I,305	218_I,306	218_I,307	218_I,308	218_I,309	218_I,310	218_I,311	218_I,312	218_I,313	218_I,314	218_I,315	218_I,316	218_I,317	218_I,318	218_I,319	218_I,320	218_I,321	218_I,322	218_I,323	218_I,324	218_I,325	218_I,326	218_I,327	218_I,328	218_I,329	218_I,330	218_I,331	218_I,332	218_I,333	218_I,334	218_I,335	218_I,336	218_I,337	218_I,338	218_I,339	218_I,340	218_I,341	218_I,342	218_I,343	218_I,344	218_I,345	218_I,346	218_I,347	218_I,348	218_I,349	218_I,350	218_I,351	218_I,352	218_I,353	218_I,354	218_I,355	218_I,356	218_I,357	218_I,358	218_I,359	218_I,360	218_I,361	218_I,362	218_I,363	218_I,364	218_I,365	218_I,366	218_I,367	218_I,368	218_I,369	218_I,370	218_I,371	218_I,372	218_I,373	218_I,374	218_I,375	218_I,376	218_I,377	218_I,378	218_I,379	218_I,380	218_I,381	218_I,382	218_I,383	218_I,384	218_I,385	218_I,386	218_I,387	218_I,388	218_I,389	218_I,390	218_I,391	218_I,392	218_I,393	218_I,394	218_I,395	218_I,396	218_I,397	218_I,398	218_I,399	218_I,400	218_I,401	218_I,402	218_I,403	218_I,404	218_I,405	218_I,406	218_I,407	218_I,408	218_I,409	218_I,410	218_I,411	218_I,412	218_I,413	218_I,414	218_I,415	218_I,416	218_I,417	218_I,418	218_I,419	218_I,420	218_I,421	218_I,422	218_I,423	218_I,424	218_I,425	218_I,426	218_I,427	218_I,428	218_I,429	218_I,430	218_I,431	218_I,432	218_I,433	218_I,434	218_I,435	218_I,436	218_I,437	218_I,438	218_I,439	218_I,440	218_I,441	218_I,442	218_I,443	218_I,444	218_I,445	218_I,446	218_I,447	218_I,448	218_I,449	218_I,450	218_I,451	218_I,452	218_I,453	218_I,454	218_I,455	218_I,456	218_I,457	218_I,458	218_I,459	218_I,460	218_I,461	218_I,462	218_I,463	218_I,464	218_I,465	218_I,466	218_I,467	218_I,468	218_I,469	218_I,470	218_I,471	218_I,472	218_I,473	218_I,474	218_I,475	218_I,476	218_I,477	218_I,478	218_I,479	218_I,480	218_I,481	218_I,482	218_I,483	218_I,484	218_I,485	218_I,486	218_I,487	218_I,488	218_I,489	218_I,490	218_I,491	218_I,492	218_I,493	218_I,494	218_I,495	218_I,496	218_I,497	218_I,498	218_I,499	218_I,500	218_I,501	218_I,502	218_I,503	218_I,504	218_I,505	218_I,506	218_I,507	218_I,508	218_I,509	218_I,510	218_I,511	218_I,512	218_I,513	218_I,514	218_I,515	218_I,516	218_I,517	218_I,518	218_I,519	218_I,520	218_I,521	218_I,522	218_I,523	218_I,524	218_I,525	218_I,526	218_I,527	218_I,528	218_I,529	218_I,530	218_I,531	218_I,532	218_I,533	218_I,534	218_I,535	218_I,536	218_I,537	218_I,538	218_I,539	218_I,540	218_I,541	218_I,542	218_I,543	218_I,544	218_I,545	218_I,546	218_I,547	218_I,548	218_I,549	218_I,550	218_I,551	218_I,552	218_I,553	218_I,554	218_I,555	218_I,556	218_I,557	218_I,558	218_I,559	218_I,560	218_I,561	218_I,562	218_I,563	218_I,564	218_I,565	218_I,566	218_I,567	218_I,568	218_I,569	218_I,570	218_I,571	218_I,572	218_I,573	218_I,574	218_I,575	218_I,576	218_I,577	218_I,578	218_I,579	218_I,580	218_I,581	218_I,582	218_I,583	218_I,584	218_I,585	218_I,586	218_I,587	218_I,588	218_I,589	218_I,590	218_I,591	218_I,592	218_I,593	218_I,594	218_I,595	218_I,596	218_I,597	218_I,598	218_I,599	218_I,600	218_I,601	218_I,602	218_I,603	218_I,604	218_I,605	218_I,606	218_I,607	218_I,608	218_I,609	218_I,610	218_I,611	218_I,612	218_I,613	218_I,614	218_I,615	218_I,616	218_I,617	218_I,618	218_I,619	218_I,620	218_I,621	218_I,622	218_I,623	218_I,624	218_I,625	218_I,626	218_I,627	218_I,628	218_I,629	218_I,630	218_I,631	218_I,632	218_I,633	218_I,634	218_I,635	218_I,636	218_I,637	218_I,638	218_I,639	218_I,640	218_I,641	218_I,642	218_I,643	218_I,644	218_I,645	218_I,646	218_I,647	218_I,648	218_I,649	218_I,650	218_I,651	218_I,652	218_I,653	218_I,654	218_I,655	218_I,656	218_I,657	218_I,658	218_I,659	218_I,660	218_I,661	218_I,662	218_I,663	218_I,664	218_I,665	218_I,666	218_I,667	218_I,668	218_I,669	218_I,670	218_I,671	218_I,672	218_I,673	218_I,674	218_I,675	218_I,676	218_I,677	218_I,678	218_I,679	218_I,680	218_I,681	218_I,682	218_I,683	218_I,684	218_I,685	218_I,686	218_I,687	218_I,688	218_I,689	218_I,690	218_I,691	218_I,692	218_I,693	218_I,694	218_I,695	218_I,696	218_I,697	218_I,698	218_I,699	218_I,700	218_I,701	218_I,702	218_I,703	218_I,704	218_I,705	218_I,706	218_I,707	218_I,708	218_I,709	218_I,710	218_I,711	218_I,712	218_I,713	218_I,714	218_I,715	218_I,716	218_I,717	218_I,718	218_I,719	218_I,720	218_I,721	218_I,722	218_I,723	218_I,724	218_I,725	218_I,726	218_I,727	218_I,728	218_I,729	218_I,730	218_I,731	218_I,732	218_I,733	218_I,734	218_I,735	218_I,736	218_I,737	218_I,738	218_I,739	218_I,740	218_I,741	218_I,742	218_I,743	218_I,744	218_I,745	218_I,746	218_I,747	218_I,748	218_I,749	218_I,750	218_I,751	218_I,752	218_I,753	218_I,754	218_I,755	218_I,756	218_I,757	218_I,758	218_I,759	218_I,760	218_I,761	218_I,762	218_I,763	218_I,764	218_I,765	218_I,766	218_I,767	218_I,768	218_I,769	218_I,770	218_I,771	218_I,772	218_I,773	218_I,774	218_I,775	218_I,776	218_I,777	218_I,778	218_I,779	218_I,780	218_I,781	218_I,782	218_I,783	218_I,784	218_I,785	218_I,786	218_I,787	218_I,788	218_I,789	218_I,790	218_I,791	218_I,792	218_I,793	218_I,794	218_I,795	218_I,796	218_I,797	218_I,798	218_I,799	218_I,800	218_I,801	218_I,802	218_I,803	218_I,804	218_I,805	218_I,806	218_I,807	218_I,808	218_I,809	218_I,810	218_I,811	218_I,812	218_I,813	218_I,814	218_I,815	218_I,816	218_I,817	218_I,818	218_I,819	218_I,820	218_I,821	218_I,822	218_I,823	218_I,824	218_I,825	218_I,826	218_I,827	218_I,828	218_I,829	218_I,830	218_I,831	218_I,832	218_I,833	218_I,834	218_I,835	218_I,836	218_I,837	218_I,838	218_I,839	218_I,840	218_I,841	218_I,842	218_I,843	218_I,844	218_I,845	218_I,846	218_I,847	218_I,848	218_I,849	218_I,850	218_I,851	218_I,852	218_I,853	218_I,854	218_I,855	218_I,856	218_I,857	218_I,858	218_I,859	218_I,860	218_I,861	218_I,862	218_I,863	218_I,864	218_I,865	218_I,866	218_I,867	218_I,868	218_I,869	218_I,870	218_I,871	218_I,872	218_I,873	218_I,874	218_I,875	218_I,876	218_I,877	218_I,878	218_I,879	218_I,880	218_I,881	218_I,882	218_I,883	218_I,884	218_I,885	218_I,886	218_I,887	218_I,888	218_I,889	218_I,890	218_I,891	218_I,892	218_I,893	218_I,894	218_I,895	218_I,896	218_I,897	218_I,898	218_I,899	218_I,900	218_I,901	218_I,902	218_I,903	218_I,904	218_I,905	218_I,906	218_I,907	218_I,908	218_I,909	218_I,910	218_I,911	218_I,912	218_I,913	218_I,914	218_I,915	218_I,916	218_I,917	218_I,918	218_I,919	218_I,920	218_I,921	218_I,922	218_I,923	218_I,924	218_I,925	218_I,926	218_I,927	218_I,928	218_I,929	218_I,930	218_I,931	218_I,932	218_I,933	218_I,934	218_I,935	218_I,936	218_I,937	218_I,938	218_I,939	218_I,940	218_I,941	218_I,942	218_I,943	218_I,944	218_I,945	218_I,946	218_I,947	218_I,948	218_I,949	218_I,950	218_I,951	218_I,952	218_I,953	218_I,954	218_I,955	218_I,956	218_I,957	218_I,958	218_I,959	218_I,960	218_I,961	218_I
---------	---------	---------	---------	---------	---------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-------

TABLE 19 (CON'T)

	216_Q_+1	216_Q_2	216_Q_1	216_U_1	216_L_+1	216_L_1	216_L_+1	216_L_1	216_V_+1	216_L_1	216_G_+1	216_F_+1	216_F_1	216_D_+1	216_D_1	216_D_+2
216_L_2	0.0594	0.077	0.0595	0.0768	0.13	0.1845	0.2392	0.2939	0.3192	0.2304	0.2111	0.0772	0.182	0.0447	0.9592	216_L_2
216_L_3	0.0724	0.093	0.0784	0.0983	0.1528	0.2248	0.3008	0.3808	0.4048	0.3748	0.9023	0.0376	0.9468	0.4577	0.8824	216_L_3
216_L_4	0.0594	0.0608	0.0595	0.0613	0.14	0.1845	0.2392	0.2939	0.3192	0.2304	0.2111	0.0772	0.182	0.0447	0.9592	216_L_4
216_L_5	0.1084	0.0988	0.0825	0.0661	0.0182	0.173	0.078	0.1059	0.0672	0.0595	0.0768	0.0376	0.9468	0.4577	0.8824	216_L_5
216_L_6	0.0693	0.0746	0.0692	0.1557	0.2158	0.3078	0.3978	0.4978	0.5178	0.4904	0.2772	0.0772	0.182	0.0447	0.9592	216_L_6
216_L_7	0.0516	0.0695	0.0537	0.062	0.1031	0.1739	0.2439	0.3139	0.3339	0.3064	0.2772	0.0772	0.182	0.0447	0.9592	216_L_7
216_L_8	0.0575	0.0597	0.0597	0.0595	0.1885	0.3351	0.4821	0.6291	0.6591	0.6316	0.3768	0.0376	0.9468	0.4577	0.8824	216_L_8
216_L_+1	0.0185	0.0358	0.0365	0.0185	0.0324	0.0305	0.0328	0.0331	0.0333	0.0335	0.0337	0.0339	0.0341	0.0343	0.0345	216_L_+1
216_L_+2	0.0388	0.0778	0.0816	0.0816	0.2332	0.3928	0.5528	0.7128	0.7328	0.7128	0.4528	0.0376	0.9468	0.4577	0.8824	216_L_+2
216_L_+4	0.0388	0.0778	0.0816	0.0816	0.2332	0.3928	0.5528	0.7128	0.7328	0.7128	0.4528	0.0376	0.9468	0.4577	0.8824	216_L_+4
216_R_+2	0.0431	0.0333	0.0365	0.0333	0.2332	0.3928	0.5528	0.7128	0.7328	0.7128	0.4528	0.0376	0.9468	0.4577	0.8824	216_R_+2
216_R_+1	0.0391	0.0291	0.0178	0.0193	0.02	0.3412	0.4162	0.4772	0.5128	0.5328	0.5128	0.2927	0.205	0.577	0.3394	216_R_+1
216_R_2	0.0168	0.0146	0.0553	0.0528	0.2428	0.3853	0.5283	0.6713	0.7128	0.7328	0.7128	0.4528	0.0376	0.9468	0.4577	216_R_2
216_R_1	0.0303	0.0168	0.0183	0.0181	0.3307	0.5853	0.8354	1.0854	1.1154	1.0854	0.8354	0.5854	0.3307	0.124	0.3672	216_R_1
216_OR_+7	0.014	0.0178	0.0816	0.08	0.0323	0.1904	0.3423	0.4943	0.5143	0.4943	0.3423	0.1904	0.124	0.3672	0.1904	216_OR_+7
216_OR_+6	0.0214	0.03082	0.0092	0.012	0.0484	0.2588	0.4682	0.6782	0.6982	0.6782	0.4682	0.2588	0.124	0.3672	0.1904	216_OR_+6
216_OR_+5	0.0251	0.0372	0.0152	0.0239	0.1068	0.4597	0.851	1.2429	1.2629	1.2429	0.851	0.4597	0.124	0.3672	0.1904	216_OR_+5
216_OR_+4	0.0083	0.00223	0.0045	0.004	0.0078	0.0551	0.3833	0.644	0.663	0.644	0.3833	0.0551	0.124	0.3672	0.1904	216_OR_+4
216_Q_+1	0.0075	0.0083	0.0083	0.0075	0.0075	0.0075	0.0075	0.0075	0.0075	0.0075	0.0075	0.0075	0.0075	0.0075	0.0075	216_Q_+1
216_Q_2	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	216_Q_2
216_Q_1	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	216_Q_1
216_U_+1	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	216_U_+1
216_L_+1	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_L_+1
216_L_+1	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_L_+1
216_L_+1	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_L_+1
216_L_+1	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_L_+1
216_V_+2	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_V_+2
216_V_+2	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_V_+2
216_V_+1	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_V_+1
216_G_+1	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_G_+1
216_F_+1	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_F_+1
216_D_+1	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_D_+1
216_D_+2	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	0.0593	216_D_+2
CNTL	58.10%	72.50%	83.40%	86.10%	87.00%	89.30%	92.20%	92.30%	92.30%	92.30%	92.30%	92.30%	92.30%	92.30%	92.30%	92.30%
CASE	50.50%	84.60%	93.40%	95.10%	97.00%	99.30%	99.50%	99.50%	99.50%	99.50%	99.50%	99.50%	99.50%	99.50%	99.50%	99.50%

TABLE 20 (CON'T)

216_Q_1	216_Q_2	216_Q_3	216_U_1	216_L_1	216_L_2	216_L_3	216_L_4	216_L_5	216_L_6	216_L_7	216_L_8	216_L_9	216_L_10	216_L_11	216_L_12	216_L_13	216_L_14	216_L_15	216_L_16	216_L_17	216_L_18	216_L_19	216_L_20	216_L_21	216_L_22	216_L_23	216_L_24	216_L_25	216_L_26	216_L_27	216_L_28	216_L_29	216_L_30	216_L_31	216_L_32	216_L_33	216_L_34	216_L_35	216_L_36	216_L_37	216_L_38	216_L_39	216_L_40	216_L_41	216_L_42	216_L_43	216_L_44	216_L_45	216_L_46	216_L_47	216_L_48	216_L_49	216_L_50	216_L_51	216_L_52	216_L_53	216_L_54	216_L_55	216_L_56	216_L_57	216_L_58	216_L_59	216_L_60	216_L_61	216_L_62	216_L_63	216_L_64	216_L_65	216_L_66	216_L_67	216_L_68	216_L_69	216_L_70	216_L_71	216_L_72	216_L_73	216_L_74	216_L_75	216_L_76	216_L_77	216_L_78	216_L_79	216_L_80	216_L_81	216_L_82	216_L_83	216_L_84	216_L_85	216_L_86	216_L_87	216_L_88	216_L_89	216_L_90	216_L_91	216_L_92	216_L_93	216_L_94	216_L_95	216_L_96	216_L_97	216_L_98	216_L_99	216_L_100	216_L_101	216_L_102	216_L_103	216_L_104	216_L_105	216_L_106	216_L_107	216_L_108	216_L_109	216_L_110	216_L_111	216_L_112	216_L_113	216_L_114	216_L_115	216_L_116	216_L_117	216_L_118	216_L_119	216_L_120	216_L_121	216_L_122	216_L_123	216_L_124	216_L_125	216_L_126	216_L_127	216_L_128	216_L_129	216_L_130	216_L_131	216_L_132	216_L_133	216_L_134	216_L_135	216_L_136	216_L_137	216_L_138	216_L_139	216_L_140	216_L_141	216_L_142	216_L_143	216_L_144	216_L_145	216_L_146	216_L_147	216_L_148	216_L_149	216_L_150	216_L_151	216_L_152	216_L_153	216_L_154	216_L_155	216_L_156	216_L_157	216_L_158	216_L_159	216_L_160	216_L_161	216_L_162	216_L_163	216_L_164	216_L_165	216_L_166	216_L_167	216_L_168	216_L_169	216_L_170	216_L_171	216_L_172	216_L_173	216_L_174	216_L_175	216_L_176	216_L_177	216_L_178	216_L_179	216_L_180	216_L_181	216_L_182	216_L_183	216_L_184	216_L_185	216_L_186	216_L_187	216_L_188	216_L_189	216_L_190	216_L_191	216_L_192	216_L_193	216_L_194	216_L_195	216_L_196	216_L_197	216_L_198	216_L_199	216_L_200	216_L_201	216_L_202	216_L_203	216_L_204	216_L_205	216_L_206	216_L_207	216_L_208	216_L_209	216_L_210	216_L_211	216_L_212	216_L_213	216_L_214	216_L_215	216_L_216	216_L_217	216_L_218	216_L_219	216_L_220	216_L_221	216_L_222	216_L_223	216_L_224	216_L_225	216_L_226	216_L_227	216_L_228	216_L_229	216_L_230	216_L_231	216_L_232	216_L_233	216_L_234	216_L_235	216_L_236	216_L_237	216_L_238	216_L_239	216_L_240	216_L_241	216_L_242	216_L_243	216_L_244	216_L_245	216_L_246	216_L_247	216_L_248	216_L_249	216_L_250	216_L_251	216_L_252	216_L_253	216_L_254	216_L_255	216_L_256	216_L_257	216_L_258	216_L_259	216_L_260	216_L_261	216_L_262	216_L_263	216_L_264	216_L_265	216_L_266	216_L_267	216_L_268	216_L_269	216_L_270	216_L_271	216_L_272	216_L_273	216_L_274	216_L_275	216_L_276	216_L_277	216_L_278	216_L_279	216_L_280	216_L_281	216_L_282	216_L_283	216_L_284	216_L_285	216_L_286	216_L_287	216_L_288	216_L_289	216_L_290	216_L_291	216_L_292	216_L_293	216_L_294	216_L_295	216_L_296	216_L_297	216_L_298	216_L_299	216_L_300	216_L_301	216_L_302	216_L_303	216_L_304	216_L_305	216_L_306	216_L_307	216_L_308	216_L_309	216_L_310	216_L_311	216_L_312	216_L_313	216_L_314	216_L_315	216_L_316	216_L_317	216_L_318	216_L_319	216_L_320	216_L_321	216_L_322	216_L_323	216_L_324	216_L_325	216_L_326	216_L_327	216_L_328	216_L_329	216_L_330	216_L_331	216_L_332	216_L_333	216_L_334	216_L_335	216_L_336	216_L_337	216_L_338	216_L_339	216_L_340	216_L_341	216_L_342	216_L_343	216_L_344	216_L_345	216_L_346	216_L_347	216_L_348	216_L_349	216_L_350	216_L_351	216_L_352	216_L_353	216_L_354	216_L_355	216_L_356	216_L_357	216_L_358	216_L_359	216_L_360	216_L_361	216_L_362	216_L_363	216_L_364	216_L_365	216_L_366	216_L_367	216_L_368	216_L_369	216_L_370	216_L_371	216_L_372	216_L_373	216_L_374	216_L_375	216_L_376	216_L_377	216_L_378	216_L_379	216_L_380	216_L_381	216_L_382	216_L_383	216_L_384	216_L_385	216_L_386	216_L_387	216_L_388	216_L_389	216_L_390	216_L_391	216_L_392	216_L_393	216_L_394	216_L_395	216_L_396	216_L_397	216_L_398	216_L_399	216_L_400	216_L_401	216_L_402	216_L_403	216_L_404	216_L_405	216_L_406	216_L_407	216_L_408	216_L_409	216_L_410	216_L_411	216_L_412	216_L_413	216_L_414	216_L_415	216_L_416	216_L_417	216_L_418	216_L_419	216_L_420	216_L_421	216_L_422	216_L_423	216_L_424	216_L_425	216_L_426	216_L_427	216_L_428	216_L_429	216_L_430	216_L_431	216_L_432	216_L_433	216_L_434	216_L_435	216_L_436	216_L_437	216_L_438	216_L_439	216_L_440	216_L_441	216_L_442	216_L_443	216_L_444	216_L_445	216_L_446	216_L_447	216_L_448	216_L_449	216_L_450	216_L_451	216_L_452	216_L_453	216_L_454	216_L_455	216_L_456	216_L_457	216_L_458	216_L_459	216_L_460	216_L_461	216_L_462	216_L_463	216_L_464	216_L_465	216_L_466	216_L_467	216_L_468	216_L_469	216_L_470	216_L_471	216_L_472	216_L_473	216_L_474	216_L_475	216_L_476	216_L_477	216_L_478	216_L_479	216_L_480	216_L_481	216_L_482	216_L_483	216_L_484	216_L_485	216_L_486	216_L_487	216_L_488	216_L_489	216_L_490	216_L_491	216_L_492	216_L_493	216_L_494	216_L_495	216_L_496	216_L_497	216_L_498	216_L_499	216_L_500	216_L_501	216_L_502	216_L_503	216_L_504	216_L_505	216_L_506	216_L_507	216_L_508	216_L_509	216_L_510	216_L_511	216_L_512	216_L_513	216_L_514	216_L_515	216_L_516	216_L_517	216_L_518	216_L_519	216_L_520	216_L_521	216_L_522	216_L_523	216_L_524	216_L_525	216_L_526	216_L_527	216_L_528	216_L_529	216_L_530	216_L_531	216_L_532	216_L_533	216_L_534	216_L_535	216_L_536	216_L_537	216_L_538	216_L_539	216_L_540	216_L_541	216_L_542	216_L_543	216_L_544	216_L_545	216_L_546	216_L_547	216_L_548	216_L_549	216_L_550	216_L_551	216_L_552	216_L_553	216_L_554	216_L_555	216_L_556	216_L_557	216_L_558	216_L_559	216_L_560	216_L_561	216_L_562	216_L_563	216_L_564	216_L_565	216_L_566	216_L_567	216_L_568	216_L_569	216_L_570	216_L_571	216_L_572	216_L_573	216_L_574	216_L_575	216_L_576	216_L_577	216_L_578	216_L_579	216_L_580	216_L_581	216_L_582	216_L_583	216_L_584	216_L_585	216_L_586	216_L_587	216_L_588	216_L_589	216_L_590	216_L_591	216_L_592	216_L_593	216_L_594	216_L_595	216_L_596	216_L_597	216_L_598	216_L_599	216_L_600	216_L_601	216_L_602	216_L_603	216_L_604	216_L_605	216_L_606	216_L_607	216_L_608	216_L_609	216_L_610	216_L_611	216_L_612	216_L_613	216_L_614	216_L_615	216_L_616	216_L_617	216_L_618	216_L_619	216_L_620	216_L_621	216_L_622	216_L_623	216_L_624	216_L_625	216_L_626	216_L_627	216_L_628	216_L_629	216_L_630	216_L_631	216_L_632	216_L_633	216_L_634	216_L_635	216_L_636	216_L_637	216_L_638	216_L_639	216_L_640	216_L_641	216_L_642	216_L_643	216_L_644	216_L_645	216_L_646	216_L_647	216_L_648	216_L_649	216_L_650	216_L_651	216_L_652	216_L_653	216_L_654	216_L_655	216_L_656	216_L_657	216_L_658	216_L_659	216_L_660	216_L_661	216_L_662	216_L_663	216_L_664	216_L_665	216_L_666	216_L_667	216_L_668	216_L_669	216_L_670	216_L_671	216_L_672	216_L_673	216_L_674	216_L_675	216_L_676	216_L_677	216_L_678	216_L_679	216_L_680	216_L_681	216_L_682	216_L_683	216_L_684	216_L_685	216_L_686	216_L_687	216_L_688	216_L_689	216_L_690	216_L_691	216_L_692	216_L_693	216_L_694	216_L_695	216_L_696	216_L_697	216_L_698	216_L_699	216_L_700	216_L_701	216_L_702	216_L_703	216_L_704	216_L_705	216_L_706	216_L_707	216_L_708	216_L_709	216_L_710	216_L_711	216_L_712	216_L_713	216_L_714	216_L_715	216_L_716	216_L_717	216_L_718	216_L_719	216_L_720	216_L_721	216_L_722	216_L_723	216_L_724	216_L_725	216_L_726	216_L_727	216_L_728	216_L_729	216_L_730	216_L_731	216_L_732	216_L_733	216_L_734	216_L_735	216_L_736	216_L_737	216_L_738	216_L_739	216_L_740	216_L_741	216_L_742	216_L_743	216_L_744	216_L_745	216_L_746	216_L_747	216_L_748	216_L_749	216_L_750	216_L_751	216_L_752	216_L_753	216_L_754	216_L_755	216_L_756	216_L_757	216_L_758	216_L_759	216_L_760	216_L_761	216_L_762	216_L_763	216_L_764	216_L_765	216_L_766	216_L_767	216_L_768	216_L_769	216_L_770	216_L_771	216_L_772	216_L_773	216_L_774	216_L_775	216_L_776	216_L_777	216_L_778	216_L_779	216_L_780	216_L_781	216_L_782	216_L_783	216_L_784	216_L_785	216_L_786	216_L_787	216_L_788	216_L_789	216_L_790	216_L_791	216_L_792	216_L_793	216_L_794	216_L_795	216_L_796	216_L_797	216_L_798	216_L_799	216_L_800	216_L_801	216_L_802	216_L_803	216_L_804	216_L_805	216_L_806	216_L_807	216_L_808	216_L_809	216_L_810	216_L_811	216_L_812	216_L_813	216_L_814	216_L_815	216_L_816	216_L_817	216_L_818	216_L_819	216_L_820	216_L_821	216_L_822	216_L_823	216_L_824	216_L_825	216_L_826	216_L_827	216_L_828	216_L_829	216_L_830	216_L_831	216_L_832	216_L_833	216_L_834	216_L_835	216_L_836	216_L_837	216_L_838	216_L_839	216_L_840	216_L_841	216_L_842	216_L_843	216_L_844	216_L_845	216_L_846	216_L_847	216_L_848	216_L_849	216_L_850	216_L_851	216_L_852	216_L_853	216_L_854	216_L_855	216_L_856	216_L_857	216_L_858	216_L_859	216_L_860	216_L_861	216_L_862	216_L_863	216_L_864	216_L_865	216_L_866	216_L_867	216_L_868	216_L_869	216_L_870	216_L_871	216_L_872	216_L_873	216_L_874	216_L_875	216_L_876	216_L_877	216_L_878	216_L_879	216_L_880	216_L_881	216_L_882	216_L_883	216_L_884	216_L_885	216_L_886	216_L_887	216_L_888	216_L_889	216_L_890	216_L_891	216_L_892	216_L_893	216_L_894	216_L_895	216_L_896	216_L_897	216_L_898	216_L_899	216_L_900	216_L_901	216_L_902	216_L_903	216_L_904	216_L_905	216_L_906	216_L_907	216_L_908	216_L_909	216_L_910	216_L_911	216_L_912	216_L_913	216_L_914	216_L_915	216_L_916	216_L_917	216_L_918	216_L_919	216_L_920	216_L_921	216_L_922	216_L_923	216_L_924	216_L_925	216_L_926	216_L_927	216_L_928	216_L_929	216_L_930	216_L_931	216_L_932	216_L_933	216_L_934	216_L_935	216_L_936	216_L_937	216_L_938	216_L_939	216_L_940	216_L_941	216_L_942	216_L_943	216_L_944	216_L_945	216_L_946	216_L_947	216_L_948	216_L_949	216_L_950	216_L_951	216_L_952	216_L_953	216_L_954	216_L_955	216_L_956	216_L_957	2
---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	---

2. Bronchial Hyper-responsiveness: A similar test for association of 2-SNP-a-time haplotypes with BHR ($PC_{20} \leq 16$ mg/ml) was performed. In Tables 21, 22 and 23, the haplotype analysis (2-at-a-time) for all SNPs in Gene 216 is presented for the combined, the UK and the US populations, respectively. One haplotype in Gene 216 (Table 21: SNPs T+2 & QR+4, $p=0.0041$) was significant at the 0.01 level in the combined sample. In contrast, in the UK population, seventeen haplotypes were significant at the 0.01 level in Gene 216 (Table 22). In the US population, nine haplotypes were significant at the 0.01 level in Gene 216 (Table 23). Tables 18, 19, and 20 and Tables 21, 22 and 23 showed similar patterns of significance with lower level achieved in the BHR analysis due to the reduced sample size in the ($PC_{20} \leq 16$ mg/ml) subgroup.

In summary, haplotype analysis of SNPs significantly strengthened the evidence in support of Gene 216 as an asthma susceptibility gene. In some SNP combinations, the association was increased by an order of magnitude. The most striking association again appeared in the 3' region of the gene, in agreement with the single SNP analysis.

TABLE 21

216 T 2	216 T 3	216 T 4	216 T 5	216 T 6	216 T 7	216 T 8	216 T 9	216 T 10	216 T 11	216 T 12	216 T 13	216 T 14	216 T 15	216 T 16	216 T 17	216 T 18	216 T 19	216 T 20	216 T 21	216 T 22	216 T 23	216 T 24	216 T 25	216 T 26	216 T 27	216 T 28	216 T 29	216 T 30	216 T 31	216 T 32	216 T 33	216 T 34	216 T 35	216 T 36	216 T 37	216 T 38	216 T 39	216 T 40	216 T 41	216 T 42	216 T 43	216 T 44	216 T 45	216 T 46	216 T 47	216 T 48	216 T 49	216 T 50	216 T 51	216 T 52	216 T 53	216 T 54	216 T 55	216 T 56	216 T 57	216 T 58	216 T 59	216 T 60	216 T 61	216 T 62	216 T 63	216 T 64	216 T 65	216 T 66	216 T 67	216 T 68	216 T 69	216 T 70	216 T 71	216 T 72	216 T 73	216 T 74	216 T 75	216 T 76	216 T 77	216 T 78	216 T 79	216 T 80	216 T 81	216 T 82	216 T 83	216 T 84	216 T 85	216 T 86	216 T 87	216 T 88	216 T 89	216 T 90	216 T 91	216 T 92	216 T 93	216 T 94	216 T 95	216 T 96	216 T 97	216 T 98	216 T 99	216 T 100	216 T 101	216 T 102	216 T 103	216 T 104	216 T 105	216 T 106	216 T 107	216 T 108	216 T 109	216 T 110	216 T 111	216 T 112	216 T 113	216 T 114	216 T 115	216 T 116	216 T 117	216 T 118	216 T 119	216 T 120	216 T 121	216 T 122	216 T 123	216 T 124	216 T 125	216 T 126	216 T 127	216 T 128	216 T 129	216 T 130	216 T 131	216 T 132	216 T 133	216 T 134	216 T 135	216 T 136	216 T 137	216 T 138	216 T 139	216 T 140	216 T 141	216 T 142	216 T 143	216 T 144	216 T 145	216 T 146	216 T 147	216 T 148	216 T 149	216 T 150	216 T 151	216 T 152	216 T 153	216 T 154	216 T 155	216 T 156	216 T 157	216 T 158	216 T 159	216 T 160	216 T 161	216 T 162	216 T 163	216 T 164	216 T 165	216 T 166	216 T 167	216 T 168	216 T 169	216 T 170	216 T 171	216 T 172	216 T 173	216 T 174	216 T 175	216 T 176	216 T 177	216 T 178	216 T 179	216 T 180	216 T 181	216 T 182	216 T 183	216 T 184	216 T 185	216 T 186	216 T 187	216 T 188	216 T 189	216 T 190	216 T 191	216 T 192	216 T 193	216 T 194	216 T 195	216 T 196	216 T 197	216 T 198	216 T 199	216 T 200	216 T 201	216 T 202	216 T 203	216 T 204	216 T 205	216 T 206	216 T 207	216 T 208	216 T 209	216 T 210	216 T 211	216 T 212	216 T 213	216 T 214	216 T 215	216 T 216	216 T 217	216 T 218	216 T 219	216 T 220	216 T 221	216 T 222	216 T 223	216 T 224	216 T 225	216 T 226	216 T 227	216 T 228	216 T 229	216 T 230	216 T 231	216 T 232	216 T 233	216 T 234	216 T 235	216 T 236	216 T 237	216 T 238	216 T 239	216 T 240	216 T 241	216 T 242	216 T 243	216 T 244	216 T 245	216 T 246	216 T 247	216 T 248	216 T 249	216 T 250	216 T 251	216 T 252	216 T 253	216 T 254	216 T 255	216 T 256	216 T 257	216 T 258	216 T 259	216 T 260	216 T 261	216 T 262	216 T 263	216 T 264	216 T 265	216 T 266	216 T 267	216 T 268	216 T 269	216 T 270	216 T 271	216 T 272	216 T 273	216 T 274	216 T 275	216 T 276	216 T 277	216 T 278	216 T 279	216 T 280	216 T 281	216 T 282	216 T 283	216 T 284	216 T 285	216 T 286	216 T 287	216 T 288	216 T 289	216 T 290	216 T 291	216 T 292	216 T 293	216 T 294	216 T 295	216 T 296	216 T 297	216 T 298	216 T 299	216 T 300	216 T 301	216 T 302	216 T 303	216 T 304	216 T 305	216 T 306	216 T 307	216 T 308	216 T 309	216 T 310	216 T 311	216 T 312	216 T 313	216 T 314	216 T 315	216 T 316	216 T 317	216 T 318	216 T 319	216 T 320	216 T 321	216 T 322	216 T 323	216 T 324	216 T 325	216 T 326	216 T 327	216 T 328	216 T 329	216 T 330	216 T 331	216 T 332	216 T 333	216 T 334	216 T 335	216 T 336	216 T 337	216 T 338	216 T 339	216 T 340	216 T 341	216 T 342	216 T 343	216 T 344	216 T 345	216 T 346	216 T 347	216 T 348	216 T 349	216 T 350	216 T 351	216 T 352	216 T 353	216 T 354	216 T 355	216 T 356	216 T 357	216 T 358	216 T 359	216 T 360	216 T 361	216 T 362	216 T 363	216 T 364	216 T 365	216 T 366	216 T 367	216 T 368	216 T 369	216 T 370	216 T 371	216 T 372	216 T 373	216 T 374	216 T 375	216 T 376	216 T 377	216 T 378	216 T 379	216 T 380	216 T 381	216 T 382	216 T 383	216 T 384	216 T 385	216 T 386	216 T 387	216 T 388	216 T 389	216 T 390	216 T 391	216 T 392	216 T 393	216 T 394	216 T 395	216 T 396	216 T 397	216 T 398	216 T 399	216 T 400	216 T 401	216 T 402	216 T 403	216 T 404	216 T 405	216 T 406	216 T 407	216 T 408	216 T 409	216 T 410	216 T 411	216 T 412	216 T 413	216 T 414	216 T 415	216 T 416	216 T 417	216 T 418	216 T 419	216 T 420	216 T 421	216 T 422	216 T 423	216 T 424	216 T 425	216 T 426	216 T 427	216 T 428	216 T 429	216 T 430	216 T 431	216 T 432	216 T 433	216 T 434	216 T 435	216 T 436	216 T 437	216 T 438	216 T 439	216 T 440	216 T 441	216 T 442	216 T 443	216 T 444	216 T 445	216 T 446	216 T 447	216 T 448	216 T 449	216 T 450	216 T 451	216 T 452	216 T 453	216 T 454	216 T 455	216 T 456	216 T 457	216 T 458	216 T 459	216 T 460	216 T 461	216 T 462	216 T 463	216 T 464	216 T 465	216 T 466	216 T 467	216 T 468	216 T 469	216 T 470	216 T 471	216 T 472	216 T 473	216 T 474	216 T 475	216 T 476	216 T 477	216 T 478	216 T 479	216 T 480	216 T 481	216 T 482	216 T 483	216 T 484	216 T 485	216 T 486	216 T 487	216 T 488	216 T 489	216 T 490	216 T 491	216 T 492	216 T 493	216 T 494	216 T 495	216 T 496	216 T 497	216 T 498	216 T 499	216 T 500	216 T 501	216 T 502	216 T 503	216 T 504	216 T 505	216 T 506	216 T 507	216 T 508	216 T 509	216 T 510	216 T 511	216 T 512	216 T 513	216 T 514	216 T 515	216 T 516	216 T 517	216 T 518	216 T 519	216 T 520	216 T 521	216 T 522	216 T 523	216 T 524	216 T 525	216 T 526	216 T 527	216 T 528	216 T 529	216 T 530	216 T 531	216 T 532	216 T 533	216 T 534	216 T 535	216 T 536	216 T 537	216 T 538	216 T 539	216 T 540	216 T 541	216 T 542	216 T 543	216 T 544	216 T 545	216 T 546	216 T 547	216 T 548	216 T 549	216 T 550	216 T 551	216 T 552	216 T 553	216 T 554	216 T 555	216 T 556	216 T 557	216 T 558	216 T 559	216 T 560	216 T 561	216 T 562	216 T 563	216 T 564	216 T 565	216 T 566	216 T 567	216 T 568	216 T 569	216 T 570	216 T 571	216 T 572	216 T 573	216 T 574	216 T 575	216 T 576	216 T 577	216 T 578	216 T 579	216 T 580	216 T 581	216 T 582	216 T 583	216 T 584	216 T 585	216 T 586	216 T 587	216 T 588	216 T 589	216 T 590	216 T 591	216 T 592	216 T 593	216 T 594	216 T 595	216 T 596	216 T 597	216 T 598	216 T 599	216 T 600	216 T 601	216 T 602	216 T 603	216 T 604	216 T 605	216 T 606	216 T 607	216 T 608	216 T 609	216 T 610	216 T 611	216 T 612	216 T 613	216 T 614	216 T 615	216 T 616	216 T 617	216 T 618	216 T 619	216 T 620	216 T 621	216 T 622	216 T 623	216 T 624	216 T 625	216 T 626	216 T 627	216 T 628	216 T 629	216 T 630	216 T 631	216 T 632	216 T 633	216 T 634	216 T 635	216 T 636	216 T 637	216 T 638	216 T 639	216 T 640	216 T 641	216 T 642	216 T 643	216 T 644	216 T 645	216 T 646	216 T 647	216 T 648	216 T 649	216 T 650	216 T 651	216 T 652	216 T 653	216 T 654	216 T 655	216 T 656	216 T 657	216 T 658	216 T 659	216 T 660	216 T 661	216 T 662	216 T 663	216 T 664	216 T 665	216 T 666	216 T 667	216 T 668	216 T 669	216 T 670	216 T 671	216 T 672	216 T 673	216 T 674	216 T 675	216 T 676	216 T 677	216 T 678	216 T 679	216 T 680	216 T 681	216 T 682	216 T 683	216 T 684	216 T 685	216 T 686	216 T 687	216 T 688	216 T 689	216 T 690	216 T 691	216 T 692	216 T 693	216 T 694	216 T 695	216 T 696	216 T 697	216 T 698	216 T 699	216 T 700	216 T 701	216 T 702	216 T 703	216 T 704	216 T 705	216 T 706	216 T 707	216 T 708	216 T 709	216 T 710	216 T 711	216 T 712	216 T 713	216 T 714	216 T 715	216 T 716	216 T 717	216 T 718	216 T 719	216 T 720	216 T 721	216 T 722	216 T 723	216 T 724	216 T 725	216 T 726	216 T 727	216 T 728	216 T 729	216 T 730	216 T 731	216 T 732	216 T 733	216 T 734	216 T 735	216 T 736	216 T 737	216 T 738	216 T 739	216 T 740	216 T 741	216 T 742	216 T 743	216 T 744	216 T 745	216 T 746	216 T 747	216 T 748	216 T 749	216 T 750	216 T 751	216 T 752	216 T 753	216 T 754	216 T 755	216 T 756	216 T 757	216 T 758	216 T 759	216 T 760	216 T 761	216 T 762	216 T 763	216 T 764	216 T 765	216 T 766	216 T 767	216 T 768	216 T 769	216 T 770	216 T 771	216 T 772	216 T 773	216 T 774	216 T 775	216 T 776	216 T 777	216 T 778	216 T 779	216 T 780	216 T 781	216 T 782	216 T 783	216 T 784	216 T 785	216 T 786	216 T 787	216 T 788	216 T 789	216 T 790	216 T 791	216 T 792	216 T 793	216 T 794	216 T 795	216 T 796	216 T 797	216 T 798	216 T 799	216 T 800	216 T 801	216 T 802	216 T 803	216 T 804	216 T 805	216 T 806	216 T 807	216 T 808	216 T 809	216 T 810	216 T 811	216 T 812	216 T 813	216 T 814	216 T 815	216 T 816	216 T 817	216 T 818	216 T 819	216 T 820	216 T 821	216 T 822	216 T 823	216 T 824	216 T 825	216 T 826	216 T 827	216 T 828	216 T 829	216 T 830	216 T 831	216 T 832	216 T 833	216 T 834	216 T 835	216 T 836	216 T 837	216 T 838	216 T 839	216 T 840	216 T 841	216 T 842	216 T 843	216 T 844	216 T 845	216 T 846	216 T 847	216 T 848	216 T 849	216 T 850	216 T 851	216 T 852	216 T 853	216 T 854	216 T 855	216 T 856	216 T 857	216 T 858	216 T 859	216 T 860	216 T 861	216 T 862	216 T 863	216 T 864	216 T 865	216 T 866	216 T 867	216 T 868	216 T 869	216 T 870	216 T 871	216 T 872	216 T 873	216 T 874	216 T 875	216 T 876	216 T 877	216 T 878	216 T 879	216 T 880	216 T 881	216 T 882	216 T 883	216 T 884	216 T 885	216 T 886	216 T 887	216 T 888	216 T 889	216 T 890	216 T 891	216 T 892	216 T 893	216 T 894	216 T 895	216 T 896	216 T 897	216 T 898	216 T 899	216 T 900	216 T 901	216 T 902	216 T 903	216 T 904	216 T 905	216 T 906	216 T 907	216 T 908	216 T 909	216 T 910	216 T 911	216 T 912	216 T 913	216 T 914	216 T 915	216 T 916	216 T 917	216 T 918	216 T 919	216 T 920	216 T 921	216 T 922	216 T 923	216 T 924	216 T 925	216 T 926	216 T 927	216 T 928	216 T 929	216 T 930	216 T 931	216 T 932	216 T 933	216 T 934	216 T 935	216 T 936	216 T 937	216 T 938	216 T 939	216 T 940	216 T 941	216 T 942	216 T 943	216 T 944	216 T 945	216 T 946	216 T 947	216 T 948	216 T 949	216 T 950	216 T 951	216 T 952	216 T 953	216 T 954	216 T 955	216 T 956	216 T 957	216 T 958	216 T 959	216 T 960	216 T 961	216 T
---------	---------	---------	---------	---------	---------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-------

TABLE 21 (CON'T)

TABLE 22

	216.I.2	216.I.3	216.I.4	216.I.5	216.I.6	216.I.7	216.I.8	216.I.+1	216.I.+2	216.I.+4	216.R.+1	216.R.+2	216.R.+3	216.QR.+1	216.QR.+2	216.QR.+3	216.QR.+4	216.QR.+5	216.QR.+6	216.QR.+7	216.QR.+8	216.QR.+9	216.QR.+10
216.I.2		0.1635	0.1946	0.4642	0.5542	0.2634	0.4757	0.4674	0.1689	0.0577	0.2683	0.3236	0.2664	0.1937	0.2869	0.41	0.151	0.0598					
216.I.3			0.3592	0.9346	0.5085	0.9737	0.9445	0.9524	0.1119	0.2317	0.5889	0.8441	0.4145	0.2901	0.4049	0.9658	0.8148	0.7043					
216.I.4				1	0.1932	0.891	0.7313	0.4727	0.1102	0.8516	0.6303	0.9016	0.2739	0.2639	0.2639	0.4975	0.7313	0.7794					
216.I.5					0.2641	0.2202	0.2221	0.2228	0.0694	0.3727	0.415	0.2666	0.1902	0.132	0.15	0.3394	0.2043	0.1956					
216.I.6					1	0.5555	0.9371	0.9371	0.2145	0.6663	0.4666	0.6354	0.3337	0.3954	0.2714	0.4571	0.346	0.0558					
216.I.7					1	1	1	0.7195	0.105	0.7668	0.5568	0.277	0.2947	0.423	0.4779	0.8161	0.7714	0.3186					
216.I.8								1	0.9889	0.6249	0.5935	0.594	0.3144	0.267	0.4123	0.9072	0.3186	0.2656					
216.I.+1								0.0454	0.0666	0.1191	0.0717	0.0153	0.0181	0.0177	0.0953	0.0403	0.0189						
216.I.+2									0.2715	0.6143	0.1610	0.1598	0.1872	0.2631	0.2774	0.0397	0.0248						
216.I.+3										0.6938	0.8941	0.4271	0.3497	0.2820	0.6728	0.5997	0.1828						
216.I.+4											0.729	0.2922	0.3083	0.3422	0.6523	0.5864	0.0463						
216.I.+5												0.1836	0.2626	0.3158	0.1576	0.2808	0.0463						
216.I.+6													0.1886	0.1984	0.1401	0.2309	0.0603						
216.I.+7														0.1041	0.1133	0.2845	0.0454						
216.I.+8															0.0702	0.2983	0.0682						
216.I.+9																0.4976	0.16						
216.I.+10																	0.0231						
216.I.+11																		0.4771					
216.I.+12																			0.1508				
216.I.+13																							
216.I.+14																							
216.I.+15																							
216.I.+16																							
216.I.+17																							
216.I.+18																							
216.I.+19																							
216.D.1																							
216.D.2																							
CNTL	65.80%	8.30%	97.10%	75.40%																			
CASE	74.00%	9.00%	59.80%	61.30%																			

TABLE 23

	216 I 2	216 I 3	216 I 4	216 I 5	216 I 6	216 I 7	216 I 8	216 I 9	216 I 10	216 I 12	216 I 14	216 I 16	216 I 18	216 I 20	216 I 22	216 I 24	216 I 26	216 I 28	216 I 30	216 I 32	216 I 34	216 I 36	216 I 38	216 I 40	216 I 42	216 I 44	216 I 46	216 I 48	216 I 50	216 I 52	216 I 54	216 I 56	216 I 58	216 I 60	216 I 62	216 I 64	216 I 66	216 I 68	216 I 70	216 I 72	216 I 74	216 I 76	216 I 78	216 I 80	216 I 82	216 I 84	216 I 86	216 I 88	216 I 90	216 I 92	216 I 94	216 I 96	216 I 98	216 I 100	216 I 102	216 I 104	216 I 106	216 I 108	216 I 110	216 I 112	216 I 114	216 I 116	216 I 118	216 I 120	216 I 122	216 I 124	216 I 126	216 I 128	216 I 130	216 I 132	216 I 134	216 I 136	216 I 138	216 I 140	216 I 142	216 I 144	216 I 146	216 I 148	216 I 150	216 I 152	216 I 154	216 I 156	216 I 158	216 I 160	216 I 162	216 I 164	216 I 166	216 I 168	216 I 170	216 I 172	216 I 174	216 I 176	216 I 178	216 I 180	216 I 182	216 I 184	216 I 186	216 I 188	216 I 190	216 I 192	216 I 194	216 I 196	216 I 198	216 I 200	216 I 202	216 I 204	216 I 206	216 I 208	216 I 210	216 I 212	216 I 214	216 I 216	216 I 218	216 I 220	216 I 222	216 I 224	216 I 226	216 I 228	216 I 230	216 I 232	216 I 234	216 I 236	216 I 238	216 I 240	216 I 242	216 I 244	216 I 246	216 I 248	216 I 250	216 I 252	216 I 254	216 I 256	216 I 258	216 I 260	216 I 262	216 I 264	216 I 266	216 I 268	216 I 270	216 I 272	216 I 274	216 I 276	216 I 278	216 I 280	216 I 282	216 I 284	216 I 286	216 I 288	216 I 290	216 I 292	216 I 294	216 I 296	216 I 298	216 I 300	216 I 302	216 I 304	216 I 306	216 I 308	216 I 310	216 I 312	216 I 314	216 I 316	216 I 318	216 I 320	216 I 322	216 I 324	216 I 326	216 I 328	216 I 330	216 I 332	216 I 334	216 I 336	216 I 338	216 I 340	216 I 342	216 I 344	216 I 346	216 I 348	216 I 350	216 I 352	216 I 354	216 I 356	216 I 358	216 I 360	216 I 362	216 I 364	216 I 366	216 I 368	216 I 370	216 I 372	216 I 374	216 I 376	216 I 378	216 I 380	216 I 382	216 I 384	216 I 386	216 I 388	216 I 390	216 I 392	216 I 394	216 I 396	216 I 398	216 I 400	216 I 402	216 I 404	216 I 406	216 I 408	216 I 410	216 I 412	216 I 414	216 I 416	216 I 418	216 I 420	216 I 422	216 I 424	216 I 426	216 I 428	216 I 430	216 I 432	216 I 434	216 I 436	216 I 438	216 I 440	216 I 442	216 I 444	216 I 446	216 I 448	216 I 450	216 I 452	216 I 454	216 I 456	216 I 458	216 I 460	216 I 462	216 I 464	216 I 466	216 I 468	216 I 470	216 I 472	216 I 474	216 I 476	216 I 478	216 I 480	216 I 482	216 I 484	216 I 486	216 I 488	216 I 490	216 I 492	216 I 494	216 I 496	216 I 498	216 I 500	216 I 502	216 I 504	216 I 506	216 I 508	216 I 510	216 I 512	216 I 514	216 I 516	216 I 518	216 I 520	216 I 522	216 I 524	216 I 526	216 I 528	216 I 530	216 I 532	216 I 534	216 I 536	216 I 538	216 I 540	216 I 542	216 I 544	216 I 546	216 I 548	216 I 550	216 I 552	216 I 554	216 I 556	216 I 558	216 I 560	216 I 562	216 I 564	216 I 566	216 I 568	216 I 570	216 I 572	216 I 574	216 I 576	216 I 578	216 I 580	216 I 582	216 I 584	216 I 586	216 I 588	216 I 590	216 I 592	216 I 594	216 I 596	216 I 598	216 I 600	216 I 602	216 I 604	216 I 606	216 I 608	216 I 610	216 I 612	216 I 614	216 I 616	216 I 618	216 I 620	216 I 622	216 I 624	216 I 626	216 I 628	216 I 630	216 I 632	216 I 634	216 I 636	216 I 638	216 I 640	216 I 642	216 I 644	216 I 646	216 I 648	216 I 650	216 I 652	216 I 654	216 I 656	216 I 658	216 I 660	216 I 662	216 I 664	216 I 666	216 I 668	216 I 670	216 I 672	216 I 674	216 I 676	216 I 678	216 I 680	216 I 682	216 I 684	216 I 686	216 I 688	216 I 690	216 I 692	216 I 694	216 I 696	216 I 698	216 I 700	216 I 702	216 I 704	216 I 706	216 I 708	216 I 710	216 I 712	216 I 714	216 I 716	216 I 718	216 I 720	216 I 722	216 I 724	216 I 726	216 I 728	216 I 730	216 I 732	216 I 734	216 I 736	216 I 738	216 I 740	216 I 742	216 I 744	216 I 746	216 I 748	216 I 750	216 I 752	216 I 754	216 I 756	216 I 758	216 I 760	216 I 762	216 I 764	216 I 766	216 I 768	216 I 770	216 I 772	216 I 774	216 I 776	216 I 778	216 I 780	216 I 782	216 I 784	216 I 786	216 I 788	216 I 790	216 I 792	216 I 794	216 I 796	216 I 798	216 I 800	216 I 802	216 I 804	216 I 806	216 I 808	216 I 810	216 I 812	216 I 814	216 I 816	216 I 818	216 I 820	216 I 822	216 I 824	216 I 826	216 I 828	216 I 830	216 I 832	216 I 834	216 I 836	216 I 838	216 I 840	216 I 842	216 I 844	216 I 846	216 I 848	216 I 850	216 I 852	216 I 854	216 I 856	216 I 858	216 I 860	216 I 862	216 I 864	216 I 866	216 I 868	216 I 870	216 I 872	216 I 874	216 I 876	216 I 878	216 I 880	216 I 882	216 I 884	216 I 886	216 I 888	216 I 890	216 I 892	216 I 894	216 I 896	216 I 898	216 I 900	216 I 902	216 I 904	216 I 906	216 I 908	216 I 910	216 I 912	216 I 914	216 I 916	216 I 918	216 I 920	216 I 922	216 I 924	216 I 926	216 I 928	216 I 930	216 I 932	216 I 934	216 I 936	216 I 938	216 I 940	216 I 942	216 I 944	216 I 946	216 I 948	216 I 950	216 I 952	216 I 954	216 I 956	216 I 958	216 I 960	216 I 962	216 I 964	216 I 966	216 I 968	216 I 970	216 I 972	216 I 974	216 I 976	216 I 978	216 I 980	216 I 982	216 I 984	216 I 986	216 I 988	216 I 990	216 I 992	216 I 994	216 I 996	216 I 998	216 I 1000	216 I 1002	216 I 1004	216 I 1006	216 I 1008	216 I 1010	216 I 1012	216 I 1014	216 I 1016	216 I 1018	216 I 1020	216 I 1022	216 I 1024	216 I 1026	216 I 1028	216 I 1030	216 I 1032	216 I 1034	216 I 1036	216 I 1038	216 I 1040	216 I 1042	216 I 1044	216 I 1046	216 I 1048	216 I 1050	216 I 1052	216 I 1054	216 I 1056	216 I 1058	216 I 1060	216 I 1062	216 I 1064	216 I 1066	216 I 1068	216 I 1070	216 I 1072	216 I 1074	216 I 1076	216 I 1078	216 I 1080	216 I 1082	216 I 1084	216 I 1086	216 I 1088	216 I 1090	216 I 1092	216 I 1094	216 I 1096	216 I 1098	216 I 1100	216 I 1102	216 I 1104	216 I 1106	216 I 1108	216 I 1110	216 I 1112	216 I 1114	216 I 1116	216 I 1118	216 I 1120	216 I 1122	216 I 1124	216 I 1126	216 I 1128	216 I 1130	216 I 1132	216 I 1134	216 I 1136	216 I 1138	216 I 1140	216 I 1142	216 I 1144	216 I 1146	216 I 1148	216 I 1150	216 I 1152	216 I 1154	216 I 1156	216 I 1158	216 I 1160	216 I 1162	216 I 1164	216 I 1166	216 I 1168	216 I 1170	216 I 1172	216 I 1174	216 I 1176	216 I 1178	216 I 1180	216 I 1182	216 I 1184	216 I 1186	216 I 1188	216 I 1190	216 I 1192	216 I 1194	216 I 1196	216 I 1198	216 I 1200	216 I 1202	216 I 1204	216 I 1206	216 I 1208	216 I 1210	216 I 1212	216 I 1214	216 I 1216	216 I 1218	216 I 1220	216 I 1222	216 I 1224	216 I 1226	216 I 1228	216 I 1230	216 I 1232	216 I 1234	216 I 1236	216 I 1238	216 I 1240	216 I 1242	216 I 1244	216 I 1246	216 I 1248	216 I 1250	216 I 1252	216 I 1254	216 I 1256	216 I 1258	216 I 1260	216 I 1262	216 I 1264	216 I 1266	216 I 1268	216 I 1270	216 I 1272	216 I 1274	216 I 1276	216 I 1278	216 I 1280	216 I 1282	216 I 1284	216 I 1286	216 I 1288	216 I 1290	216 I 1292	216 I 1294	216 I 1296	216 I 1298	216 I 1300	216 I 1302	216 I 1304	216 I 1306	216 I 1308	216 I 1310	216 I 1312	216 I 1314	216 I 1316	216 I 1318	216 I 1320	216 I 1322	216 I 1324	216 I 1326	216 I 1328	216 I 1330	216 I 1332	216 I 1334	216 I 1336	216 I 1338	216 I 1340	216 I 1342	216 I 1344	216 I 1346	216 I 1348	216 I 1350	216 I 1352	216 I 1354	216 I 1356	216 I 1358	216 I 1360	216 I 1362	216 I 1364	216 I 1366	216 I 1368	216 I 1370	216 I 1372	216 I 1374	216 I 1376	216 I 1378	216 I 1380	216 I 1382	216 I 1384	216 I 1386	216 I 1388	216 I 1390	216 I 1392	216 I 1394	216 I 1396	216 I 1398	216 I 1400	216 I 1402	216 I 1404	216 I 1406	216 I 1408	216 I 1410	216 I 1412	216 I 1414	216 I 1416	216 I 1418	216 I 1420	216 I 1422	216 I 1424	216 I 1426	216 I 1428	216 I 1430	216 I 1432	216 I 1434	216 I 1436	216 I 1438	216 I 1440	216 I 1442	216 I 1444	216 I 1446	216 I 1448	216 I 1450	216 I 1452	216 I 1454	216 I 1456	216 I 1458	216 I 1460	216 I 1462	216 I 1464	216 I 1466	216 I 1468	216 I 1470	216 I 1472	216 I 1474	216 I 1476	216 I 1478	216 I 1480	216 I 1482	216 I 1484	216 I 1486	216 I 1488	216 I 1490	216 I 1492	216 I 1494	216 I 1496	216 I 1498	216 I 1500	216 I 1502	216 I 1504	216 I 1506	216 I 1508	216 I 1510	216 I 1512	216 I 1514	216 I 1516	216 I 1518	216 I 1520	216 I 1522	216 I 1524	216 I 1526	216 I 1528	216 I 1530	216 I 1532	216 I 1534	216 I 1536	216 I 1538	216 I 1540	216 I 1542	216 I 1544	216 I 1546	216 I 1548	216 I 1550	216 I 1552	216 I 1554	216 I 1556	216 I 1558	216 I 1560	216 I 1562	216 I 1564	216 I 1566	216 I 1568	216 I 1570	216 I 1572	216 I 1574	216 I 1576	216 I 1578	216 I 1580	216 I 1582	216 I 1584	216 I 1586	216 I 1588	216 I 1590	216 I 1592	216 I 1594	216 I 1596	216 I 1598	216 I 1600	216 I 1602	216 I 1604	216 I 1606	216 I 1608	216 I 1610	216 I 1612	216 I 1614	216 I 1616	216 I 1618	216 I 1620	216 I 1622	216 I 1624	216 I 1626	216 I 1628	216 I 1630	216 I 1632	216 I 1634	216 I 1636	216 I 1638	216 I 1640	216 I 1642	216 I 1644	216 I 1646	216 I 1648	216 I 1650	216 I 1652	216 I 1654	216 I 1656	216 I 1658	216 I 1660	216 I 1662	216 I 1664	216 I 1666	216 I 1668	216 I 1670	216 I 1672	216 I 1674	216 I 1676	216 I 1678	216 I 1680	216 I 1682	216 I 1684	216 I 1686	216 I 1688	216 I 1690	216 I 1692	216 I 1694	216 I 1696	216 I 1698	216 I 1700	216 I 1702	216 I 1704	216 I 1706	216 I 1708	216 I 1710	216 I 1712	216 I 1714	216 I 1716	216 I 1718	216 I 1720	216 I 1722	216 I 1724	216 I 1726	216 I 1728	216 I 1730	216 I 1732	216 I 1734	216 I 1736	216 I 1738	216 I 1740	216 I 1742	216 I 1744	216 I 1746	216 I 1748	216 I 1750	216 I 1752	216 I 1754	216 I 1756	216 I 1758	216 I 1760	216 I 1762	216 I 1764	216 I 1766	216 I 1768	216 I 1770	216 I 1772	216 I 1774	216 I 1776	216 I 1778	216 I 1780	216 I 1782	216 I 1784	216 I 1786	216 I 1788	216 I 1790	216 I 1792	216 I 1794	216 I 1796	216 I 1798	216 I 1800	216 I 1802	216 I 1804	216 I 1806	216 I 1808	216 I 1810	216 I 1812	216 I 1814	216 I 1816	216 I 1818	216 I 1820	216 I 1822	216 I 1824	216 I 1826	216 I 1828	216 I 1830	216 I 1832	216 I 1834	216 I 1836	216 I 1838	216 I 1840	21
--	---------	---------	---------	---------	---------	---------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	----

TABLE 23 (CON'T)

216_Q_+1	216_Q_2	216_Q_1	216_U_-1	216_U_-1	216_L_-1	216_L_-1	216_L_-2	216_V_-2	216_V_+1	216_L_+1	216_Q_+1	216_Q_+1	216_F_+1	216_F_+1	216_D_1	216_D_1	216_D_2	216_D_2
0.1389	0.1647	0.2214	0.1925	0.0531	0.0638	0.0588	0.0918	0.1041	0.1021	0.0503	0.1687	0.0815	0.0815	0.0482	0.0482	0.0362	216_I_2	
0.7352	0.1073	0.68	0.9027	0.1168	0.9596	0.9112	0.9818	0.9519	0.8271	0.1768	0.9577	0.1256	0.1256	0.0815	0.0815	0.0362	216_I_3	
0.218_+2	0.0918	0.6201	0.4625	0.4541	0.0516	0.0516	0.9347	0.2813	0.3793	0.0455	0.3741	0.0476	0.0476	0.9752	0.9752	0.0362	216_I_4	
0.6441	0.1575	0.6062	0.1038	0.0531	0.0638	0.0588	0.9819	0.8548	0.1185	0.0652	0.2079	0.0943	0.0943	0.9568	0.9568	0.0362	216_I_5	
0.3381	0.0814	0.9135	0.8378	0.0579	0.9759	0.9759	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_I_6	
0.8372	0.037	0.8639	0.0475	0.0753	0.9759	0.9759	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_I_7	
0.6355	0.0261	0.8372	0.4935	0.0753	0.9759	0.9759	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_I_8	
0.4503	0.046	0.519	0.365	0.055	0.055	0.055	0.9877	0.784	0.8165	0.7192	0.3821	0.7548	0.7548	0.9568	0.9568	0.0362	216_I_9	
0.3669	0.0971	0.8588	0.8408	0.0711	0.9877	0.9877	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_I_10	
0.5765	0.0192	0.7132	0.655	0.055	0.055	0.055	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_I_11	
0.1575	0.0789	0.5406	0.4359	0.055	0.055	0.055	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_I_12	
0.1518	0.0838	0.1592	0.1183	0.055	0.055	0.055	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_I_13	
0.1628	0.0846	0.1592	0.1183	0.055	0.055	0.055	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_I_14	
0.216_R_+2	0.0112	0.0275	0.0075	0.0075	0.0075	0.0075	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_R_+1	
0.216_R_+1	0.0112	0.0275	0.0075	0.0075	0.0075	0.0075	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_R_+2	
0.216_R_+2	0.0112	0.0275	0.0075	0.0075	0.0075	0.0075	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_R_+3	
0.216_R_+3	0.0112	0.0275	0.0075	0.0075	0.0075	0.0075	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_R_+4	
0.216_R_+4	0.0112	0.0275	0.0075	0.0075	0.0075	0.0075	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_R_+5	
0.216_Q_+1	0.0112	0.0275	0.0075	0.0075	0.0075	0.0075	0.9877	0.8435	0.835	0.0677	0.957	0.957	0.957	0.9568	0.9568	0.0362	216_Q_+2	
0.216_Q_+2	0.0112	0.0275																

EXAMPLE 14: Transmission Disequilibrium Test (TDT)

To ensure that the significant association observed in the case-control studies was not an artifact due to population admixture, a family based test of association, the transmission disequilibrium test (TDT) was conducted. By
5 selecting a single affected offspring in each family, the TDT test performed a test of association (due to linkage disequilibrium) in the presence of linkage. The test determined whether a particular allele was preferentially transmitted to an affected individual over what would be expected by chance. Only heterozygous parents were considered informative for the TDT. In addition, to
10 increase power, heterozygous parents transmitting a different allele to two affected offspring were ignored. Accordingly, the TDT would be based on the same families that contributed to the linkage signal. The significance levels were estimated by Markov Chain Monte Carlo simulation methods as implemented in TDTEX from the S.A.G.E. program (Department of
15 Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, MetroHealth Campus, Case Western Reserve University, Cleveland, OH (1997)).

1. Asthma Phenotype: Five candidate SNPs were typed in the extended population in order to confirm the association seen in the case-control study. The five SNPs were in Gene 216 exons T5, T8, T+1, R1, and Q1. Since only heterozygote parents contribute information to the TDT test, SNP haplotypes (all 2-at-a-time and all 3-at-a-time) were constructed based on family data with the program GENEHUNTER (Kruglyak et al., 1996) in addition to analyzing the SNPs separately. This served to increase the informativeness
20 of the single SNPs. These haplotypes were then used as "alleles" in future TDT analyses. In addition, p-values obtained from the TDT analyses were compared to the p-values obtained from the haplotyping in the case/control setting. To check for consistency, the p-values were recorded to compare the haplotype frequencies between the cases and controls of the over-transmitted
25 alleles/haplotypes.
30

The TDT results strongly supported the association previously observed

in the case control studies (Table 24). Three of the five SNPs showed alleles that were preferentially transmitted to affected offspring ($p < 0.04$ to < 0.0044) in either the combined or UK population. When these SNPs were haplotyped together, most combinations had a haplotype that was preferentially transmitted to affected offspring ($p < 0.03$ to < 0.001). The most significant haplotype in the combined population was composed of SNPs T+1/R1/Q1 ($p = 0.0006$). The most significant haplotype in the UK population was composed of SNPs T5/R1/Q1 ($p = 0.0005$). In contrast to the UK population, none of the single SNP allele or multiple SNP haplotypes were preferentially over-transmitted to affected offspring at significant levels in the US population. This is most likely due to the combination of reduced power of the TDT versus the case-control study and the smaller sample size in the US.

Importantly, for all of the single SNP or multiple SNP haplotypes the allele that was significantly over-transmitted in either the combined population or in the UK sample was more frequent in the cases than in the controls. A summary of the TDT analyses and a comparison between the Case/control and TDT results are presented in Table 24.

2. Bronchial Hyper-responsiveness: The TDT analyses were repeated using only those asthmatic pairs that satisfied the additional criteria of having a $PC_{20} \leq 16$ mg/ml (Table 25). The vast majority of single SNP and multiple SNP haplotypes showed increased significance with the more restricted phenotype. P values reached levels of < 0.00008 for T5/R1/Q1 in the combined population and $p < 0.000008$ in the UK sample. Similar to the yes/no phenotype, for the majority of the alleles in both the combined and UK population, the over-transmitted alleles in the TDT were more frequent in the cases. Similar to the yes/no phenotype with the less powerful TDT test, no significant results were observed with smaller US sample. In summary, the analysis of single SNPs and SNP haplotypes by the TDT test provided confirmatory evidence for Gene 216 as an asthma susceptibility gene.

TABLE 24

Asthma Yes/NO				
Combined US and UK				
Over-Transmitted Haplotype				
Exon in Gene 216	TDT p-value	Case/Control p-value	Control Frequency	Case Frequency
Q 1	0.0337	0.0213	89.5%	94.8%
R 1	0.0725	NS	88.7%	88.2%
T +1	0.0956	0.0055	85.2%	92.4%
T 8	1.0000	NS	NA	
T 5	0.1364	0.0420	76.7%	83.3%
R1Q1	0.0042	0.1362	78.2%	83.1%
T+1Q1	0.0932	0.0049	85.2%	92.4%
T8Q1	0.0553	0.0084	86.0%	92.9%
T5Q1	0.2659	0.0342	76.2%	83.0%
T+1R1	0.0029	0.0465	73.9%	80.6%
T8R1	0.0799	NS	85.1%	67.9%
T5R1	0.0107	0.1537	66.1%	71.5%
T8T+1	0.2762	0.0044	85.2%	92.4%
T5T+1	0.3078	0.0012	72.5%	83.0%
T5T8	0.0948	0.0028	73.7%	83.4%
T+1R1Q1	0.0006	0.0430	73.9%	80.8%
T8R1Q1	0.0086	0.0552	74.7%	81.2%
T5R1Q1	0.0025	0.1591	65.9%	71.2%
T5T+1R1	0.0136	0.0175	62.3%	71.2%
T8T+1R1	0.0084	0.0377	73.9%	80.9%
T5T8R1	0.0060	0.0235	63.0%	71.5%
T5T8Q1	0.1242	0.0033	73.1%	83.0%
T5T8T+1	0.1540	0.0009	72.7%	83.0%
T8T+1Q1	0.1351	0.0043	85.3%	92.4%
T5T+1Q1	0.1080	0.0010	72.5%	83.0%

NS = non-significant or over-transmitted allele not present more often in cases than controls
 NA = no alleles were over-transmitted

TABLE 24 (CON'T)

Asthma Yes/NO				
UK				
Exon in Gene 216	TDT p-value	Over-Transmitted Haplotype		Case Frequency
		Case/Control p-value	Control Frequency	
Q 1	0.0044	0.0274	89.4%	95.1%
R 1	0.3665	0.1473	86.8%	91.4%
T +1	0.0128	0.0105	86.4%	93.8%
T 8	1.0000	NS	NA	
T 5	0.0434	0.0426	75.4%	83.3%
R1Q1	0.0044	0.0069	76.2%	86.5%
T+1Q1	0.0714	0.0066	86.4%	93.8%
T8Q1	0.0342	0.0275	87.4%	93.6%
T5Q1	0.1687	0.0314	74.9%	82.9%
T+1R1	0.0269	0.0018	73.2%	85.1%
T8R1	0.4848	0.0933	84.6%	89.9%
T5R1	0.0639	0.0067	63.1%	74.7%
T8T+1	0.2254	0.0069	86.4%	93.8%
T5T+1	0.2007	0.0088	72.9%	82.9%
T5T8	0.0277	0.0103	73.7%	83.4%
T+1R1Q1	0.0063	0.0016	73.2%	85.1%
T8R1Q1	0.0139	0.0039	74.1%	85.0%
T5R1Q1	0.0005	0.0136	63.4%	74.2%
T5T+1R1	0.0220	0.0036	61.5%	74.2%
T8T+1R1	0.0043	0.0012	73.2%	85.1%
T5T8R1	0.0095	0.0018	61.5%	74.7%
T5T8Q1	0.0074	0.0105	73.3%	82.9%
T5T8T+1	0.0255	0.0082	73.0%	82.9%
T8T+1Q1	0.0207	0.0087	86.4%	93.8%
T5T+1Q1	0.0127	0.0093	72.9%	82.9%

NS = non-significant or over-transmitted allele not present more often in cases than controls;
 NA = no alleles were over-transmitted

TABLE 24 (CON'T)

Asthma Yes/NO				
US				
Over-Transmitted Haplotype				
Exon in Gene 216	TDT p-value	Case/Control p-value	Control Frequency	Case Frequency
Q 1	0.8039	NS	10.4%	6.3%
R 1	0.1067	NS	92.2%	75.9%
T +1	0.6288	NS	17.1%	13.0%
T 8	1.0000	NS	NA	
T 5	0.7020	NS	20.8%	16.7%
R1Q1	0.2134	NS	81.8%	69.6%
T+1Q1	0.6811	NS	10.4%	9.7%
T8Q1	0.7584	0.2887	83.6%	90.2%
T5Q1	0.8284	NS	9.7%	8.3%
T+1R1	0.0658	NS	75.1%	63.0%
T8R1	0.0687	NS	86.1%	72.2%
T5R1	0.1859	NS	71.4%	59.3%
T8T+1	0.9465	0.4778	83.0%	87.0%
T5T+1	0.8537	0.5074	9.7%	13.0%
T5T8	0.8848	NS	20.8%	13.0%
T+1R1Q1	0.1569	NS	75.2%	62.7%
T8R1Q1	0.2386	NS	75.8%	66.0%
T5R1Q1	0.0831	NS	70.7%	59.3%
T5T+1R1	0.1332	NS	64.1%	59.9%
T8T+1R1	0.1299	NS	75.2%	63.4%
T5T8R1	0.0813	NS	65.5%	60.2%
T5T8Q1	0.8654	NS	9.7%	7.8%
T5T8T+1	0.8546	NS	9.6%	9.3%
T8T+1Q1	0.6864	NS	10.4%	9.3%
T5T+1Q1	0.8618	0.9991	9.7%	9.7%

NS = non-significant or over-transmitted allele not present more often in cases than controls;
 NA = no alleles were over-transmitted

TABLE 25

BHR				
Combined US and UK				
Exon in Gene 216	TDT p-value	Over-Transmitted Haplotype		
		Case/Control p-value	Control Frequency	Case Frequency
Q 1	0.0800	0.1565	89.5%	94.2%
R 1	0.0374	NS	88.7%	88.3%
T +1	0.1252	0.1413	85.2%	90.6%
T 8	1.0000	NS	NA	
T 5	0.0947	0.4681	76.7%	80.2%
R1Q1	0.0017	0.2040	78.2%	83.7%
T+1Q1	0.1835	0.1192	85.2%	90.6%
T8Q1	0.1616	0.0987	86.0%	91.8%
T5Q1	0.1496	0.3214	76.2%	80.2%
T+1R1	0.0015	0.1479	73.9%	80.2%
T8R1	0.0281	0.7994	85.1%	85.9%
T5R1	0.0009	0.6419	66.1%	68.4%
T8T+1	0.6224	0.1380	85.2%	90.6%
T5T+1	0.4821	0.0660	72.5%	80.3%
T5T8	0.1786	0.1284	73.7%	80.2%
T+1R1Q1	0.0003	0.1426	73.9%	80.4%
T8R1Q1	0.0035	0.1298	74.7%	81.4%
T5R1Q1	0.0001	0.4524	65.9%	69.7%
T5T+1R1	0.0052	0.1332	62.3%	69.6%
T8T+1R1	0.0066	0.1397	73.9%	80.6%
T5T8R1	0.0028	0.2632	63.0%	68.4%
T5T8Q1	0.3680	0.0954	73.1%	80.3%
T5T8T+1	0.5282	0.0786	72.7%	80.3%
T8T+1Q1	0.3105	0.1261	85.3%	90.6%
T5T+1Q1	0.5276	0.0686	72.5%	80.3%

NS = non-significant or over-transmitted allele not present more often in cases than controls;
 NA = no alleles were over-transmitted

TABLE 25 (CON'T)

BHR				
UK				
		Over-Transmitted Haplotype		
Exon in Gene	TDT p-value	Case/Control p-value	Control Frequency	Case Frequency
216				
Q 1	0.0069	0.0613	89.4%	95.8%
R 1	0.3285	0.1041	86.8%	93.0%
T +1	0.0201	0.0454	86.4%	94.0%
T 8	1.0000	NS	NA	
T 5	0.0367	0.2644	75.4%	81.6%
R1Q1	0.00078	0.0052	76.2%	89.8%
T+1Q1	0.0209	0.0280	86.4%	94.0%
T8Q1	0.0120	0.0933	87.4%	93.8%
T5Q1	0.0974	0.1624	74.9%	81.7%
T+1R1	0.0001	0.0026	73.2%	87.6%
T8R1	0.2818	0.1182	84.6%	91.0%
T5R1	0.0038	0.0420	63.1%	74.6%
T8T+1	0.1437	0.0327	86.4%	94.0%
T5T+1	0.0902	0.0739	72.9%	81.7%
T5T8	0.0536	0.1052	73.7%	81.7%
T+1R1Q1	0.000075	0.0042	73.2%	87.8%
T8R1Q1	0.0031	0.0056	74.1%	87.7%
T5R1Q1	0.0000078	0.0331	63.4%	75.4%
T5T+1R1	0.0071	0.0131	61.5%	75.3%
T8T+1R1	0.0023	0.0034	73.2%	87.8%
T5T8R1	0.0073	0.0216	61.5%	74.6%
T5T8Q1	0.0424	0.0835	73.3%	81.7%
T5T8T+1	0.1380	0.0761	73.0%	81.7%
T8T+1Q1	0.0322	0.0319	86.4%	94.0%
T5T+1Q1	0.1096	0.0756	72.9%	81.7%

NS = non-significant or over-transmitted allele not present more often in cases than controls;
 NA = no alleles were over-transmitted

TABLE 25 (CON'T)

BHR

US

Exon in Gene 216	TDT p-value	Over-Transmitted Haplotype		Control Frequency	Case Frequency
		Case/Control p-value			
Q_1	0.5081	0.7250		10.4%	12.5%
R_1	0.0577	NS		92.2%	71.4%
T_+1	0.5493	0.5937		17.1%	21.4%
T_8	1.0000	NS		NA	
T_5	0.7741	0.6206		20.8%	25.0%
R1Q1	0.1259	NS		81.8%	58.8%
T+1Q1	0.7495	0.1224		10.4%	21.4%
T8Q1	0.7514	0.7864		10.4%	12.1%
T5Q1	0.1029	0.1408		9.7%	18.8%
T+1R1	0.2012	NS		75.1%	50.0%
T8R1	0.0880	NS		86.1%	67.9%
T5R1	0.0963	NS		71.4%	46.4%
T8T+1	0.7557	0.2626		10.7%	17.9%
T5T+1	0.4904	0.0908		9.7%	21.4%
T5T8	0.8871	0.9876		20.8%	21.4%
T+1R1Q1	0.0828	NS		75.2%	50.0%
T8R1Q1	0.1759	NS		75.8%	55.9%
T5R1Q1	0.2046	NS		70.7%	46.4%
T5T+1R1	0.1915	NS		64.1%	46.4%
T8T+1R1	0.2537	NS		75.2%	50.0%
T5T8R1	0.1633	NS		65.5%	46.4%
T5T8Q1	0.6920	0.3863		9.7%	16.1%
T5T8T+1	0.8586	0.3158		9.6%	17.9%
T8T+1Q1	0.7517	0.3367		10.4%	17.9%
T5T+1Q1	0.8579	0.1166		9.7%	21.4%

NS = non-significant or over-transmitted allele not present more often in cases than controls;

NA = no alleles were over-transmitted

EXAMPLE 15: Attributable Risk Assessment

From the knowledge of the frequency of a functional polymorphism and the relative risk of the heterozygote and homozygote (at-risk) genotypes, one can evaluate the attributable fraction (M.J. Khoury et al., 1993, Fundamentals of Genetic Epidemiology, J.L. Kelsey et al., (eds), *Monographs in Epidemiology and Biostatistics*, Oxford University Press, New York, NY, Section 3, pp 74-77) or attributable risk in the population. An attributable fraction of 25% would mean that if the population were monomorphic for the protective allele, the prevalence of the trait would be 25% lower.

The formula for the attributable fraction is:

$$\text{Attributable fraction} = \frac{(1-f)^2 + 2f(1-f)\gamma + f^2\eta - 1}{(1-f)^2 + 2f(1-f)\gamma + f^2\eta},$$

where f is the allele frequency, γ is the relative risk of the heterozygote genotype over the wild type homozygote, and η is the risk of the homozygote mutant over the wild type homozygote. This approach requires the estimation of f , γ and η . Ideally these quantities should be estimated in an epidemiological sample.

The study design (genome scan with affected sibling pairs followed by association study using IBD = 2 individuals as cases in the case/control comparison) offers maximum power to detect linkage and association, but does not provide estimates of the required parameters, namely 1) the relative risk (or odds ratio) of the genotype/allele for most SNPs or haplotypes and 2) the frequency of the SNP in the general population. In a recent paper, Altshuler et al. used the data from a TDT analysis to estimate allele and genotype relative risks assuming a multiplicative model or $\eta = \gamma^2$ (D. Altshuler et al., 2000, *Nature Genetics* 26:76-80). Thus, the mutant homozygote is predicted to carry a relative risk equal to the square of the risk for the heterozygote.

To overcome some of the difficulties mentioned above that are associated with a case/control design, the data obtained from typing 5 SNPs in Gene 216 on the entire population (not just the subset of IBD = 2 individuals)

were used to estimate the relative risk of these 5 SNPs. The data from the TDT obtained by using the first asthmatic sibling per family were used. Because of the limited number of informative matings in the TDT analysis, a multiplicative model for the genotype relative risk was used as in the Altshuler et. al paper, i.e. $\eta = \gamma^2$. An interval on the attributable fraction estimates was made by constructing individual confidence regions for the allele frequency in the control population and for the attributable risk obtained from the TDT data. While combining these two confidence intervals to obtain a confidence region for the attributable fraction did not lead to a proper confidence region with the required coverage, it determined the variability involved in estimating the attributable fraction. As a short hand notation, this is referred to as a confidence interval with coverage equal to the one used for the constituent parameters.

By using the control population to estimate allele frequencies, the attributable risk was underestimated. Based on these assumptions, the attributable risk for the single SNPs that were significant in the case-control study ($p < 0.05$) in either population was computed. The AF was also computed for all SNP combinations significant in the combined TDT analysis ($p < 0.01$) using the asthma phenotype. These values are shown below.

20	SNP(s)	Attributable fraction (AF) estimate	80% Confidence Interval
	Q 1	50%	17 to 65%
	R 1	37%	4 to 57%
	T + 1	39%	7 to 57%
	T 5	22%	0 to 35%
25	R1 Q1	36%	14 to 54%
	T +1 R1	29%	8 to 47%
	T +1 R1 Q1	34%	14 to 52%
	T 5 R1 Q1	19%	3 to 38%
	T 5 T 8 R1	24%	9 to 41%
30	T 8 R1 Q1	32%	11 to 50%
	T 8 T+1 R1	25%	2 to 44%

Because the alleles that confer increased risk of developing asthma are so common (haplotype frequencies ranging from 60% to 83%), their effect translated into a substantial population attributable risk, with estimates ranging from 19 to 50% for different SNPs or SNP haplotypes. These computations
5 depended heavily on allele frequency and risk estimates. Proper estimates of the attributable fraction are based on a population sample and are only meaningful for functional SNPs or SNP haplotypes.

Conclusion: Gene 216 has been demonstrated to be an asthma gene in accordance with the data disclosed herein, including: 1) localization to a
10 region on chromosome 20 identified through linkage; 2) polymorphism analysis performed to identify sequence variants localized in the candidate gene; 3) genotype analyses of the identified polymorphisms; 4) association between identified alleles and the asthma phenotype in a case-control analysis; 5) association between identified alleles and the asthma phenotype in
15 transmission disequilibrium tests (TDT), haplotype analyses, and analyses using additional phenotypes; 6) identification of transcripts in tissues relevant to pulmonary disease and/or inflammation; and 7) characterization of Gene 216 as an ADAM family member. In addition to respiratory diseases, Gene 216 is likely to be involved in obesity and inflammatory bowel disease, as
20 obesity (Wilson et al., 1999, *Arch. Intern. Med.* **159**: 2513-14) and inflammatory bowel disease (B. Wallaert et al., 1995, *J. Exp. Med.* **182**:1897-1904) have been linked to asthma.

EXAMPLE 16: Protein Expression And Purification

Expression and purification of the Gene 216 protein of the invention can
25 be performed essentially as outlined below. To facilitate the cloning, expression, and purification of membrane and secreted protein from the 20p13-p12, a gene expression system, such as the pET System (Novagen), for cloning and expression of recombinant proteins in *E. coli* is selected. Also, a DNA sequence encoding a peptide tag, the His-Tag, is fused to the 3' end of
30 DNA sequences of interest to facilitate purification of the recombinant protein products. The 3' end is selected for fusion to avoid alteration of any 5' terminal

signal sequence.

Nucleic acids chosen, for example, from the nucleic acids set forth in SEQ ID NO:1 or SEQ ID NO:6 (Figures 24 and 29, respectively) for cloning the genes are prepared by polymerase chain reaction (PCR). Synthetic
5 oligonucleotide primers specific for the 5' and 3' ends of the nucleotide sequences are designed and purchased from Life Technologies. All forward primers (specific for the 5' end of the sequence) are designed to include an *NcoI* cloning site at the 5' terminus. These primers are designed to permit initiation of protein translation at the methionine residue encoded within the
10 *NcoI* site followed by a valine residue and the protein encoded by the DNA sequence. All reverse primers (specific for the 3' end of the sequence) include an *EcoRI* site at the 5' terminus to permit cloning of the sequence into the reading frame of the pET-28b. The pET-28b vector provides a sequence encoding an additional 20 carboxyl-terminal amino acids including six histidine
15 residues (at the C-terminus), which comprise the histidine affinity tag.

DNA prepared from the 20p13-p12 region is used as the source of template DNA for PCR amplification (Ausubel et al., 1994). To amplify a DNA sequence containing the nucleotide sequence, c DNA (50 ng) is introduced into a reaction vial containing 2 mM $MgCl_2$, 1 μM synthetic oligonucleotide primers
20 (forward and reverse primers) complementary to and flanking a defined 20p13-p12 region, 0.2 mM of each of deoxynucleotide triphosphate, dATP, dGTP, dCTP, dTTP and 2.5 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ) in a final volume of 100 microliters.

Upon completion of thermal cycling reactions, each sample of amplified
25 DNA is purified using the Qiaquick Spin PCR purification kit. All amplified DNA samples are subjected to digestion with the restriction endonucleases, e.g., *NcoI* and *EcoRI* (NEB) (Ausubel et al., 1994). DNA samples are then subjected to electrophoresis on 1.0% NuSeive (FMC BioProducts) agarose gels. DNA is visualized by exposure to ethidium bromide and long wave UV
30 irradiation. DNA contained in slices isolated from the agarose gel was purified using the BIO 101 GeneClean Kit protocol.

The pET-28b vector is prepared for cloning by digestion with restriction endonucleases, e.g., *Nco*I and *Eco*RI (NEB) (Ausubel et al., 1994). The pET-28a vector, which encodes the histidine affinity tag that can be fused to the 5' end of an inserted gene, is prepared by digestion with appropriate restriction endonucleases.

Following digestion, DNA inserts are cloned (Ausubel et al., 1994) into the previously digested pET-28b expression vector. Products of the ligation reaction are then used to transform the BL21 strain of *E. coli* (Ausubel et al., 1994) as described below.

Competent bacteria, *E. coli* strain BL21 or *E. coli* strain BL21 (DE3), are transformed with recombinant pET expression plasmids carrying the cloned sequence according to standard methods (Ausubel et al., 1994). Briefly, 1 microliter of ligation reaction is mixed with 50 microliters of electrocompetent cells and subjected to a high voltage pulse, after which samples were incubated in 0.45 ml SOC medium (0.5% yeast extract, 2.0% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ and 20 mM glucose) at 37°C with shaking for 1 hr. Samples are then spread on LB agar plates containing 25 µg/ml kanamycin sulfate for growth overnight. Transformed colonies of BL21 are then picked and analyzed to evaluate cloned inserts, as described below.

Individual BL21 clones transformed with recombinant pET-28b. 20p13-p12 region nucleotide sequences are analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers specific for the 20p13-p12 region sequences that are used in the original PCR amplification cloning reactions. Successful amplification verifies the integration of the sequence in the expression vector (Ausubel et al., 1994).

Individual clones of recombinant pET-28b vectors carrying properly cloned 20p13-p12 region nucleotide sequences are picked and incubated in 5 ml of LB broth plus 25 µg/ml kanamycin sulfate overnight. The following day plasmid DNA is isolated and purified using the QIAGEN plasmid purification protocol.

The pET vector can be propagated in any *E. coli* K-12 strain, e.g., HMS174, HB101, JM109, DH5, and the like, for purposes of cloning or plasmid preparation. Hosts for expression include *E. coli* strains containing a chromosomal copy of the gene for T7 RNA polymerase. These hosts are
5 lysogens of bacteriophage DE3, a lambda derivative that carries the *lacI* gene, the *lacUV5* promoter, and the gene for T7 RNA polymerase. T7 RNA polymerase is induced by addition of isopropyl- β -D-thiogalactoside (IPTG), and the T7 RNA polymerase transcribes any target plasmid containing a functional T7 promoter, such as pET-28b, carrying its gene of interest. Strains include,
10 for example, BL21(DE3) (Studier et al., 1990, *Meth. Enzymol.*, **185**:60-89).

To express the recombinant sequence, 50 ng of plasmid DNA are isolated as described above to transform competent BL21(DE3) bacteria as described above (provided by Novagen as part of the pET expression kit). The *lacZ* gene (β -galactosidase) is expressed in the pET-System as described for
15 the 20p13-p12 region recombinant constructions. Transformed cells were cultured in SOC medium for 1 hr, and the culture is then plated on LB plates containing 25 μ g/ml kanamycin sulfate. The following day, the bacterial colonies are pooled and grown in LB medium containing kanamycin sulfate (25 μ g/ml) to an optical density at 600 nm of 0.5 to 1.0 OD units, at which point 1
20 mM IPTG was added to the culture for 3 hr to induce gene expression of the 20p13-p12 region recombinant DNA constructions.

After induction of gene expression with IPTG, bacteria are collected by centrifugation in a Sorvall RC-3B centrifuge at 3500 x g for 15 min at 4°C. Pellets are resuspended in 50 ml of cold mM Tris-HCl, pH 8.0, 0.1 M NaCl and
25 0.1 mM EDTA (STE buffer). Cells are then centrifuged at 2000 x g for 20 min at 4°C. Wet pellets are weighed and frozen at -80°C until ready for protein purification.

A variety of methodologies known in the art can be used to purify the isolated proteins (Coligan et al., 1995, *Current Protocols in Protein Science*,
30 John Wiley & Sons, New York, NY). For example, the frozen cells can be thawed, resuspended in buffer, and ruptured by several passages through a

small volume microfluidizer (Model M-110S, Microfluidics International Corp., Newton, MA). The resultant homogenate is centrifuged to yield a clear supernatant (crude extract) and, following filtration, the crude extract is fractionated over columns. Fractions are monitored by absorbance at OD₂₈₀ nm and peak fractions may be analyzed by SDS-PAGE.

The concentrations of purified protein preparations are quantified spectrophotometrically using absorbance coefficients calculated from amino acid content (Perkins, 1986, *Eur. J. Biochem.*, **157**:169-180). Protein concentrations are also measured by the method of Bradford, 1976, *Anal. Biochem.*, **72**:248-254; and Lowry et al., 1951, *J. Biol. Chem.*, **193**:265-275 using bovine serum albumin as a standard.

SDS-polyacrylamide gels of various concentrations are purchased from Bio-Rad, and stained with Coomassie blue. Molecular weight markers may include rabbit skeletal muscle myosin (200 kDa), *E. coli* β -galactosidase (116 kDa), rabbit muscle phosphorylase B (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), egg white lysozyme (14.4 kDa) and bovine aprotinin (6.5 kDa).

Proteins can also be isolated by other conventional means of protein biochemistry and purification to obtain a substantially pure product, i.e., 80, 95, or 99% free of cell component contaminants, as described in Jacoby, 1984, *Methods in Enzymology*, Vol. 104, Academic Press, NY; Scoopes, 1987, *Protein Purification, Principles and Practice*, 2nd Ed., Springer-Verlag, NY; and Deutscher (ed), 1990, *Guide to Protein Purification, Methods in Enzymology*, Vol. 182. If the protein is secreted, it can be isolated from the supernatant in which the host cell is grown; otherwise, it can be isolated from a lysate of the host cells.

Once a sufficient quantity of the desired protein has been obtained, it may be used for various purposes. One use of the protein or polypeptide is the production of antibodies specific for binding. These antibodies may be either polyclonal or monoclonal, and may be produced by *in vitro* or *in vivo*

techniques well known in the art. Monoclonal antibodies to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas (Kohler, 1975, *Nature*, **256**:495). In summary, a mouse is inoculated with a few micrograms of protein over a period of 2 weeks. The mouse is then sacrificed. The cells that produce antibodies are then removed from the mouse's spleen. The spleen cells are then fused with polyethylene glycol with mouse myeloma cells. The successfully fused cells are diluted in a microtiter plate and growth of the culture is continued. The amount of antibody per well is measured by immunoassay methods such as ELISA (Engvall, 1980, *Meth. Enzymol.*, **70**:419). Clones producing antibody can be expanded and further propagated to produce protein antibodies. Other suitable techniques involve *in vitro* exposure of lymphocytes to the antigenic polypeptides, or alternatively, to selection of libraries of antibodies in phage or similar vectors. See Huse et al., 1989, *Science*, **246**:1275-1281. For additional information on antibody production see Davis et al., 1989, *Basic Methods in Molecular Biology*, Elsevier, NY, Section 21-2. Such antibodies are particularly useful in diagnostic assays for detection of variant protein forms, or as an active ingredient in a pharmaceutical composition.

The disclosure of each of the patents, patent applications, and publications cited in the specification is hereby incorporated by reference herein in its entirety.

Although the invention has been set forth in detail, one skilled in the art will recognize that numerous changes and modifications can be made, and that such changes and modifications may be made without departing from the spirit and scope of the invention.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:
 - 5 a. SEQ ID NO:1;
 - b. a nucleotide sequence encoding amino acid SEQ ID NO:4;
 - c. a nucleotide sequence complementary to SEQ ID NO:1;
 - d. a nucleotide sequence which hybridizes under high stringency conditions to SEQ ID NO:1;
 - 10 e. a nucleotide sequence which hybridizes under moderate stringency conditions to SEQ ID NO:1;
 - f. a nucleotide sequence which hybridizes under low stringency conditions to SEQ ID NO:1;
 - g. a nucleotide sequence which is at least 95% identical to the sequence of SEQ ID NO:1;
 - 15 h. a nucleotide sequence which is at least 80% identical to the sequence of SEQ ID NO:1; and
 - i. a nucleotide sequence which is at least 50% identical to the sequence of SEQ ID NO:1.
- 20 2. The isolated nucleic acid of claim 1 which is DNA.
3. The isolated nucleic acid of claim 1 which is RNA.
4. A vector comprising the isolated nucleic acid of claim 1.
5. A host cell comprising the expression vector of claim 4.
6. The host cell of claim 5 which is selected from the group consisting of eukaryotic and prokaryotic cells.
- 25 7. The host cell of claim 5 which is selected from the group consisting of bacterial, fungal.
8. The isolated nucleic acid of claim 1, wherein the nucleic acid

sequence comprises at least 50 consecutive nucleotides.

9. A vector comprising the isolated nucleic acid of claim 8.
10. A host cell comprising the expression vector of claim 9.
11. The host cell of claim 10 which is selected from the group
5 consisting of eukaryotic and prokaryotic cells.
12. The host cell of claim 10 which is selected from the group
consisting of bacterial, yeast, insect, mammalian, and plant cells.
13. The isolated nucleic acid of claim 1, wherein the nucleic acid
sequence comprises at least 15 consecutive nucleotides.
- 10 14. A vector comprising the isolated nucleic acid of claim 13.
15. A host cell comprising the vector of claim 14.
16. The host cell of claim 15 which is selected from the group
consisting of eukaryotic and prokaryotic cells.
17. The host cell of claim 15 which is selected from the group
15 consisting of bacterial, yeast, insect, mammalian, and plant cells.
18. An isolated nucleic acid variant which comprises the sequence
of SEQ ID NO:6, and contains at least one single nucleotide polymorphism set
forth in Table 10.
19. An isolated nucleic acid variant which comprises at least 50
20 consecutive nucleotides of SEQ ID NO:6, and contains at least one single
nucleotide polymorphism set forth in Table 10.
20. An isolated nucleic acid variant which comprises at least 15
consecutive nucleotides of SEQ ID NO:6, and contains at least one single
nucleotide polymorphism set forth in Table 10.
- 25 21. The isolated nucleic acid variant of claim 20, wherein the single
nucleotide polymorphism is selected from the group consisting of T4, T5, T8,

T+1, T+2, R1, Q1, Q2, QR+4, QR+6, QR+7, and U-1.

22. The isolated nucleic acid variant of claim 20, wherein the single nucleotide polymorphism selected from the group consisting of D1, F1, I1, L1, R2, T6, T1, T2, T3, and T7.

5 23. The isolated nucleic acid variant of claim 20 containing at least two single nucleotide polymorphisms selected from the group consisting of:

- a. T+2 and QR+4;
- b. QR+5 and QR+4;
- c. QR+4 and Q+1;
- 10 d. QR+6 and Q2; and
- e. QR+4 and Q2.

24. The isolated nucleic acid variant of claim 20, wherein the single nucleotide polymorphism is selected from the group consisting of:

- a. T5 and T8;
- 15 b. T+2 and QR+4;
- c. T4 and T5.
- d. T+1 and R1 and Q1; and
- e. T5 and R1 and Q1.

25. An isolated nucleic acid variant which comprises the sequence
20 of SEQ ID NO:1, and contains at least one single nucleotide polymorphism at a site shown in Figure 24.

26. An isolated nucleic acid variant which comprises at least 50 consecutive nucleotides of SEQ ID NO:1, and contains at least one single nucleotide polymorphism at a site shown in Figure 24.

25 27. An isolated nucleic acid variant which comprises at least 15 consecutive nucleotides of SEQ ID NO:1, and contains at least one single nucleotide polymorphism at a site shown in Figure 24.

28. An isolated alternate splice variant which comprises at least one

exon of SEQ ID NO:1 set forth in Figures 9 and 10.

29. An isolated alternate splice variant which comprises at least one exon of SEQ ID NO:1 selected from the group consisting of exons T, R, Q, and U set forth in Figures 9 and 10.

5 30. An isolated alternate splice variant which comprises at least one exon of SEQ ID NO:1 selected from the group consisting of exons A, B, C, D, D', E, F, G, H, I, J, K, L, L2, M, N, O, P, and S set forth in Figures 9 and 10.

31. An isolated alternate splice variant which comprises a sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:350-362.

10 32. An isolated polypeptide encoded by the nucleic acid of any one of claims 1 and 8.

33. An isolated polypeptide encoded by the nucleic acid of any one of claims 18, 19, 25, and 26.

15 34. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

- a. SEQ ID NO:4;
- b. an amino acid sequence which is at least 80% identical to SEQ ID NO:4;
- c. an amino acid sequence which is at least 75% identical to SEQ ID NO:4; and
- 20 d. an amino acid sequence which is at least 65% identical to SEQ ID NO:4.

35. An isolated polypeptide comprising at least 20 consecutive residues of the amino acid sequence of claim 34.

25 36. An isolated polypeptide comprising at least 7 consecutive residues of the amino acid sequence of claim 34.

37. An antibody or antibody fragment which binds to the isolated

polypeptide of claim 32.

38. An antibody or antibody fragment which binds to the isolated polypeptide of claim 33.

39. An antibody or antibody fragment which binds to the isolated
5 polypeptide according to any one of claims 34-36

40. The antibody or antibody fragment of claim 37 which is selected from the group consisting of polyclonal and monoclonal antibodies.

41. The antibody or antibody fragment of claim 38 which is selected from the group consisting of polyclonal and monoclonal antibodies.

10 42. The antibody or antibody fragment of claim 39 which is selected from the group consisting of polyclonal and monoclonal antibodies.

43. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

- 15 a. SEQ ID NO:6;
b. a nucleotide sequence comprising at least 50 consecutive nucleotides of SEQ ID NO:6; and
c. a nucleotide sequence comprising at least 15 consecutive nucleotides of SEQ ID NO:6.

20 44. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

- a. SEQ ID NO:364;
b. a nucleotide sequence complementary to SEQ ID NO:364.
c. a nucleotide sequence comprising at least 50 consecutive nucleotides of SEQ ID NO:364;
25 d. a nucleotide sequence comprising at least 15 consecutive nucleotides of SEQ ID NO:364.
e. SEQ ID NO:365;
f. a nucleotide sequence complementary to SEQ ID NO:365;

- g. a nucleotide sequence comprising at least 50 consecutive nucleotides of SEQ ID NO:365; and
 - h. a nucleotide sequence comprising at least 15 consecutive nucleotides of SEQ ID NO:365.
- 5 45. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
- a. SEQ ID NO:366;
 - b. an amino acid sequence comprising 20 consecutive residues of SEQ ID NO:366; and
 - 10 c. an amino acid sequence comprising 7 consecutive residues of SEQ ID NO:M366.
46. An isolated antibody or antibody fragment that binds to the isolated polypeptide of claim 45.
47. The antibody or antibody fragment of claim 46 which is selected
15 from the group consisting of monoclonal and polyclonal antibodies.
48. An isolated antisense nucleic acid comprising at least 15 consecutive nucleotides of a sequence complementary to SEQ ID NO:1.
49. An isolated antisense nucleic acid comprising at least 15 consecutive nucleotides of a sequence complementary to SEQ ID NO:6.
- 20 50. A vector comprising the isolated antisense nucleic acid of any one of claims 48-49.
51. A kit for detecting a Gene 216 nucleotide sequence comprising:
- a. the isolated nucleic acid of any one of claims 13, 20, and 27; and
 - b. at least one component to detect binding of the isolated nucleic
25 acid to a Gene 216 nucleotide sequence.
52. A kit for detecting a Gene 216 amino acid sequence comprising:
- a. the isolated antibody of claim 42; and
 - b. at least one component to detect binding of the isolated antibody

to a Gene 216 amino acid sequence.

53. A method of identifying a Gene 216 ligand, comprising:

- a. contacting the isolated polypeptide of claim 35 with a test agent under conditions that allow the polypeptide to bind to the test agent, and
5 thereby form a complex; and
- b. detecting the polypeptide-test agent complex of (a), wherein detection of the complex indicates identification of a Gene 216 ligand.

54. The method of claim 53, wherein the ligand is a metalloprotease inhibitor.

10 55. The method of claim 54, wherein the metalloprotease inhibitor is a proglutamyl peptide analog.

56. The method of claim 55, wherein the proglutamyl peptide analog is an analog of pyroGlu-Asn-Trp-OH or pyroGlu-Glu-Trp-OH.

15 57. A pharmaceutical composition comprising the ligand identified according to the method of any one of claims 53-56, and a physiologically acceptable carrier, excipient, or diluent.

58. A pharmaceutical composition comprising the isolated nucleic acid of any one of claims 1, 8, 13, 43, 48, and 49, and a physiologically acceptable carrier, excipient, or diluent.

20 59. A pharmaceutical composition comprising the vector of any one of claims 4, 9, 14, and 48, and a physiologically acceptable carrier, excipient, or diluent.

25 60. A pharmaceutical composition comprising the isolated antibody or antibody fragment of claim 42, and a physiologically acceptable carrier, excipient, or diluent.

61. A pharmaceutical composition comprising the isolated polypeptide of claim 36 and a physiologically acceptable carrier, excipient, or

diluent.

62. A method of identifying a human Gene 216 or ortholog, comprising:

- 5 a. contacting the nucleic acid of any one of claims 1, 8, and 13 with a biological sample under conditions that allow the nucleic acid to hybridize to a nucleic acid in the sample, and thereby form a complex; and
- b. detecting the hybridization complex of (a), wherein detection of the complex indicates identification of a human Gene 216 or ortholog.

63. A method of treating a chromosome 20 disorder comprising
10 administering the pharmaceutical composition of claim 57 in an amount effective to treat the disorder.

64. The method of claim 63, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory bowel disease.

15 65. A method of treating a chromosome 20 disorder comprising administering the pharmaceutical composition of claim 58 in an amount effective to treat the disorder.

20 66. The method of claim 65, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory bowel disease.

67. A method of treating a chromosome 20 disorder comprising administering the pharmaceutical composition of claim 59 in an amount effective to treat the disorder.

25 68. The method of claim 67, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory bowel disease.

69. A method of treating a chromosome 20 disorder comprising administering the pharmaceutical composition of claim 60 in an amount

effective to treat the disorder.

70. The method of claim 69, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory bowel disease.

5 71. A method of treating a chromosome 20 disorder comprising administering the pharmaceutical composition of claim 61 in an amount effective to treat the disorder.

72. The method of claim 71, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory
10 bowel disease.

73. A transgenic mouse whose genome comprises an introduced null mutation in an endogenous Gene 216.

74. The transgenic mouse of claim 73, wherein both alleles of the endogenous Gene 216 of said mouse have been disrupted.

15 75. The transgenic mouse of claim 74, wherein the mouse genome further comprises a human Gene 216 nucleic acid sequence.

76. A method of making a homozygous transgenic knockout mouse comprising:

- 20 a. disrupting an endogenous Gene 216 in mouse embryonic stem cells;
- b. introducing said embryonic stem cells into a mouse blastocyst and transplanting said blastocyst into a pseudopregnant mouse;
- c. allowing said blastocyst to develop into a chimeric mouse;
- d. breeding said chimeric mouse to produce offspring; and
- 25 e. screening said offspring to identify a homozygous transgenic knockout mouse.

77. A method of making a knockout mouse comprising administering the antibody or antibody fragment of claim 47 in an amount effective to disrupt

endogenous Gene 216 polypeptide function, thereby making a knockout mouse.

78. A method of forming a crystal of the isolated Gene 216 polypeptide of claim 36 comprising:

- 5 a. incubating the polypeptide with a solution selected from the group consisting of the solutions in wells 1-30 in Table 1 under conditions to allow crystalization; and
- b. detecting the crystalization in (a), whereby crystalization indicates formation of a Gene 216 polypeptide crystal.

10 79. A method of diagnosing a chromosome 20 disorder, comprising:
 a. contacting the isolated nucleic acid of any one of claims 20-24 with a biological sample under high stringency conditions that allow the nucleic acid to hybridize to a nucleic acid in the sample, and thereby form a complex; and

15 b. detecting the hybridization complex of (a), wherein detection of the complex indicates diagnosis of a chromosome disorder.

80. The method of claim 79, wherein the disorder is selected from the group consisting of asthma, obesity, and inflammatory bowel disease.

20 81. A method of diagnosing a chromosome 20 disorder comprising:
 a. contacting the isolated antibody or antibody fragment of claim 41 with a biological sample under high stringency conditions that allow the antibody or antibody fragment to bind to an amino acid sequence in the sample, and thereby form a complex; and

25 b. detecting the complex of (a), wherein detection of the complex indicates diagnosis of a chromosome disorder.

82. A method of determining a pharmacogenetic profile comprising:

- a. contacting the isolated nucleic acid of any one of claims 20-24 with a biological sample under high stringency conditions that allow the nucleic acid to hybridize to a nucleic acid in the sample, and thereby form a complex;

and

b. detecting the hybridization complex of (a), wherein detection of the complex determines the pharmacogenetic profile.

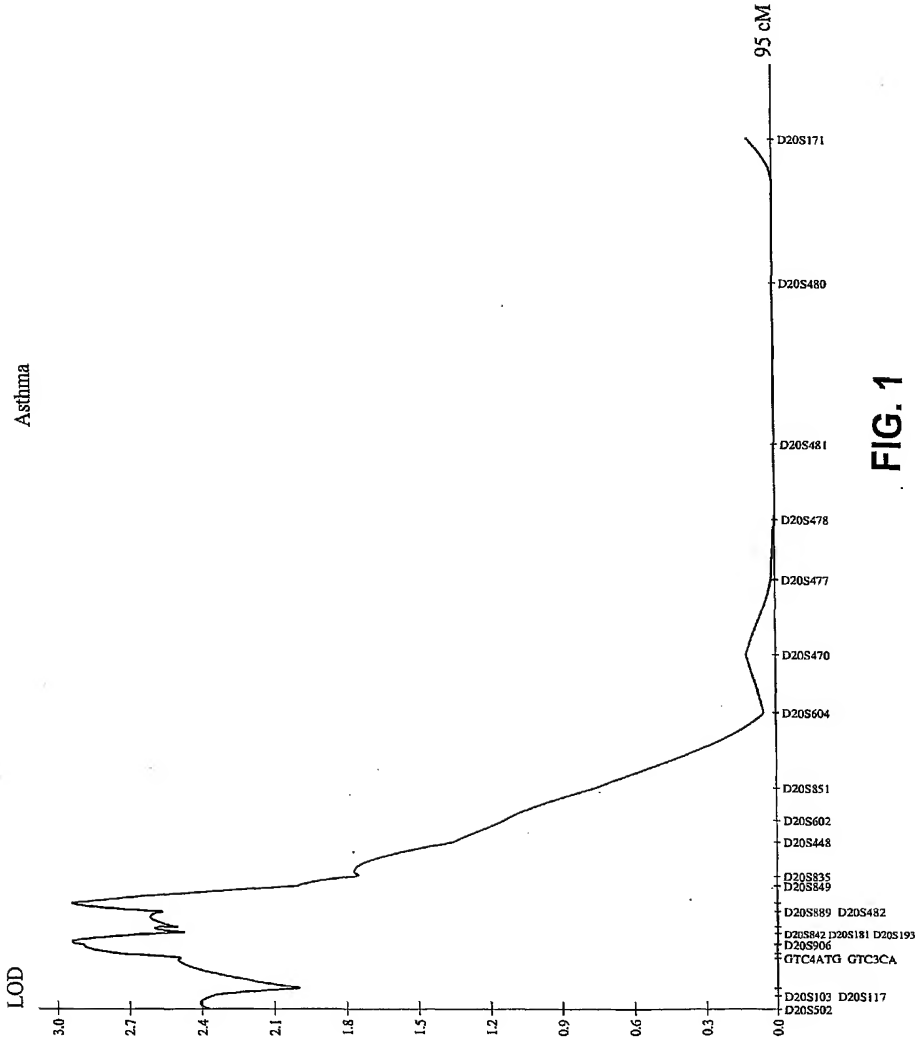
83. A method of determining a pharmacogenetic profile comprising:

5 a. contacting the isolated antibody of claim 41 with a biological sample under high stringency conditions that allow the antibody to hybridize to an amino acid sequence in the sample, and thereby form a complex; and

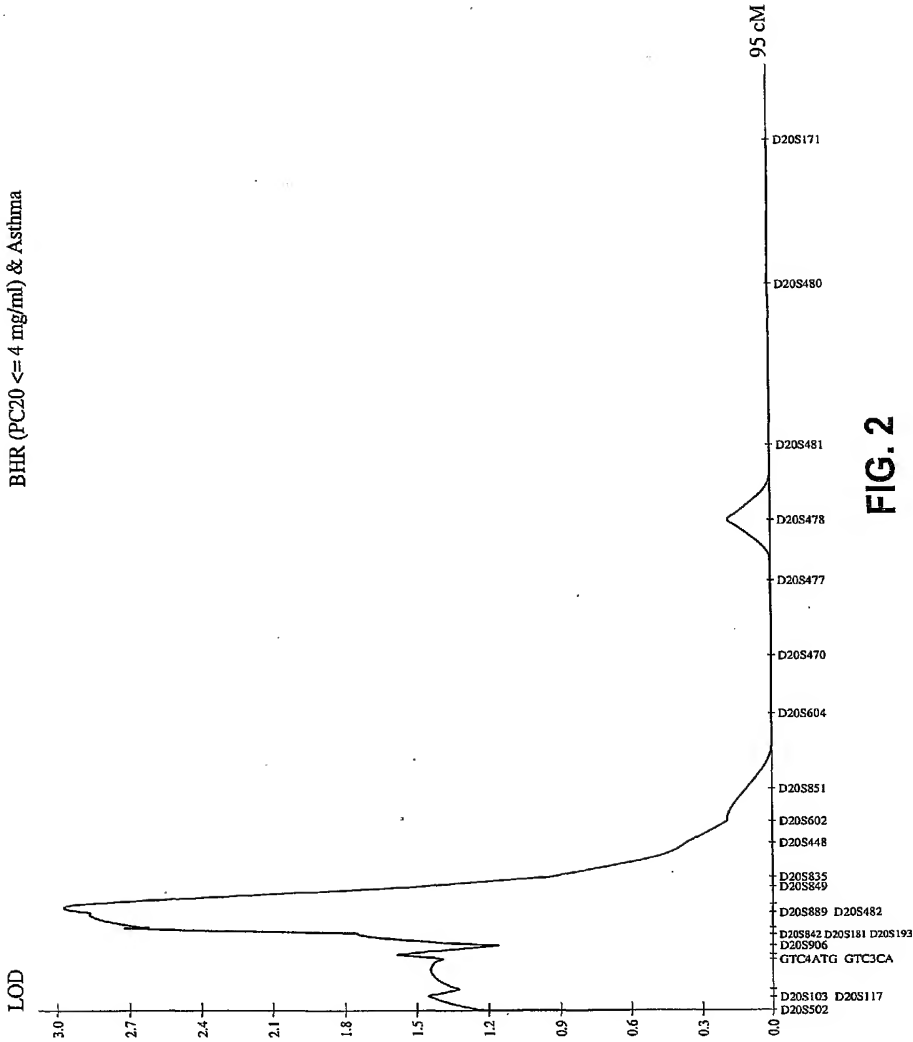
b. detecting the complex of (a), wherein detection of the complex determines the pharmacogenetic profile.

10 84. A cell line comprising the isolated nucleic acid of any one of claims 8, 19, 26, and 28.

85. A biochip comprising the isolated nucleic acid of any one of claims 8, 19, 26, and 28.



BHR (PC20 <= 4 mg/ml) & Asthma



BHR (PC20 \leq 16 mg/ml) & Asthma

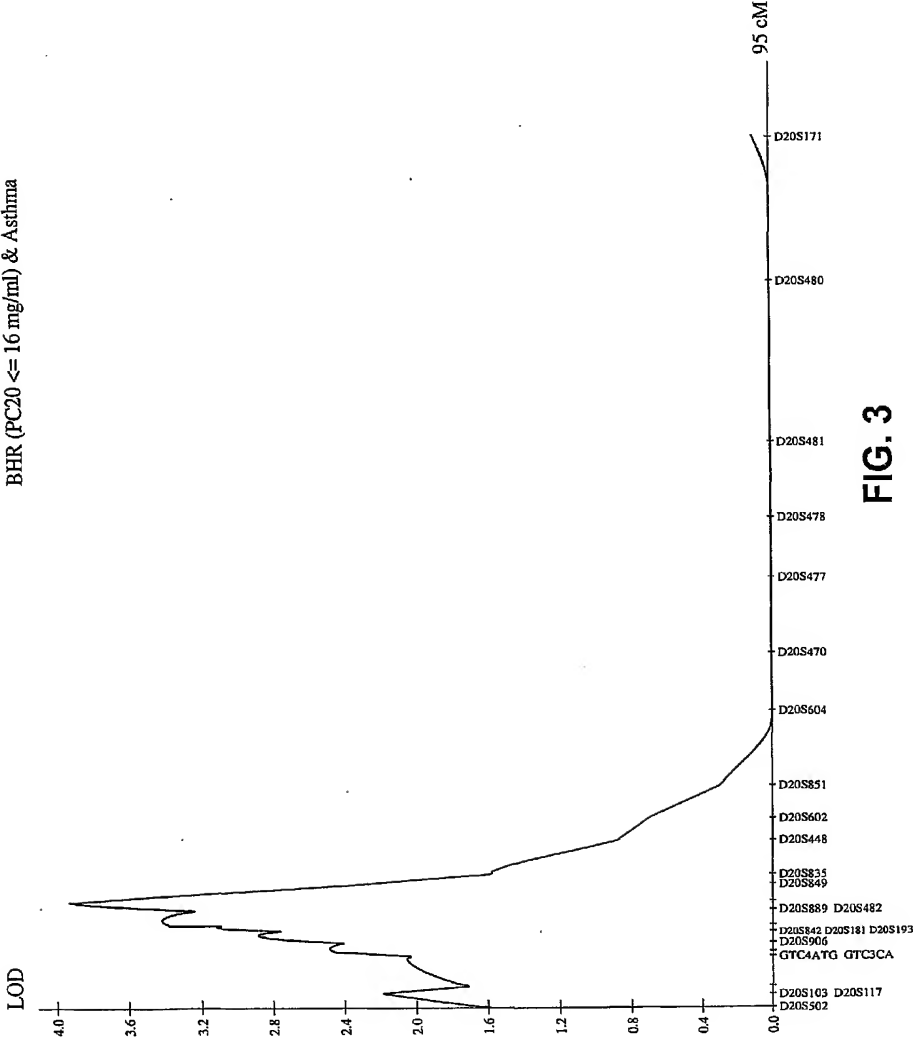


FIG. 3

High Total IgE & Asthma

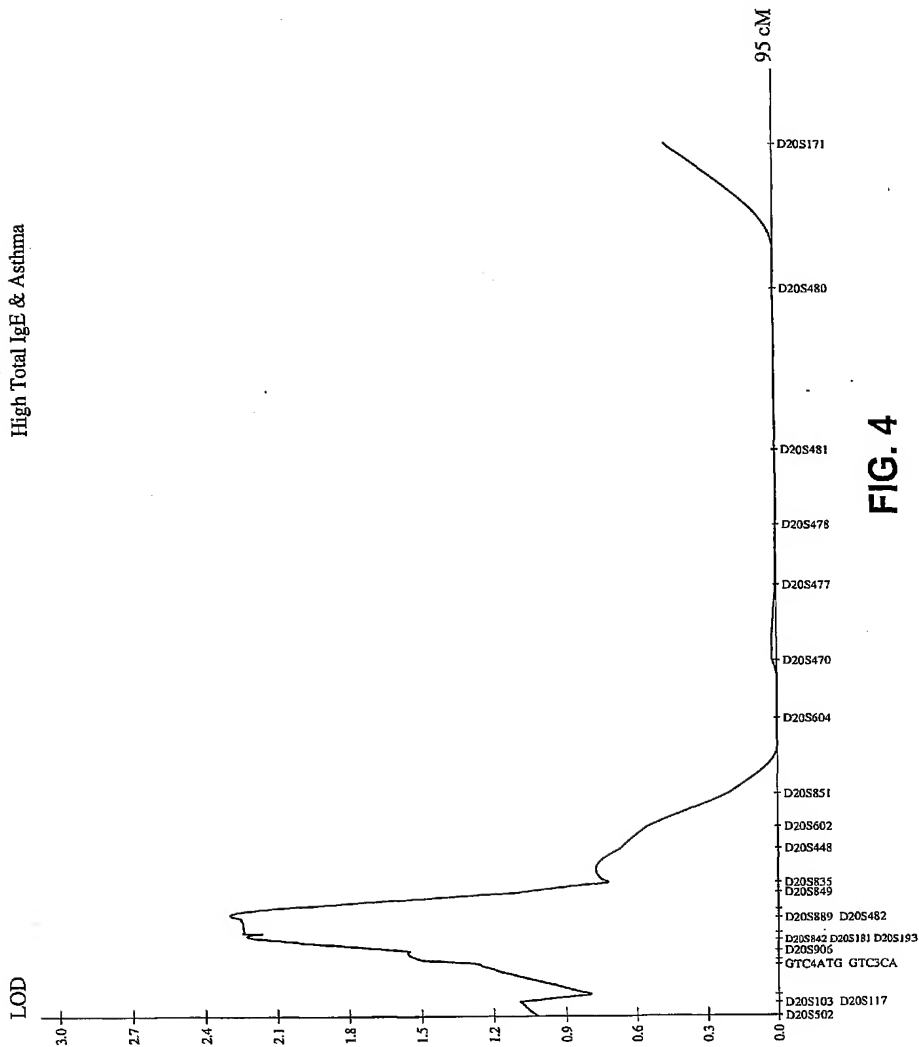


FIG. 4

High Specific IgE & Asthma

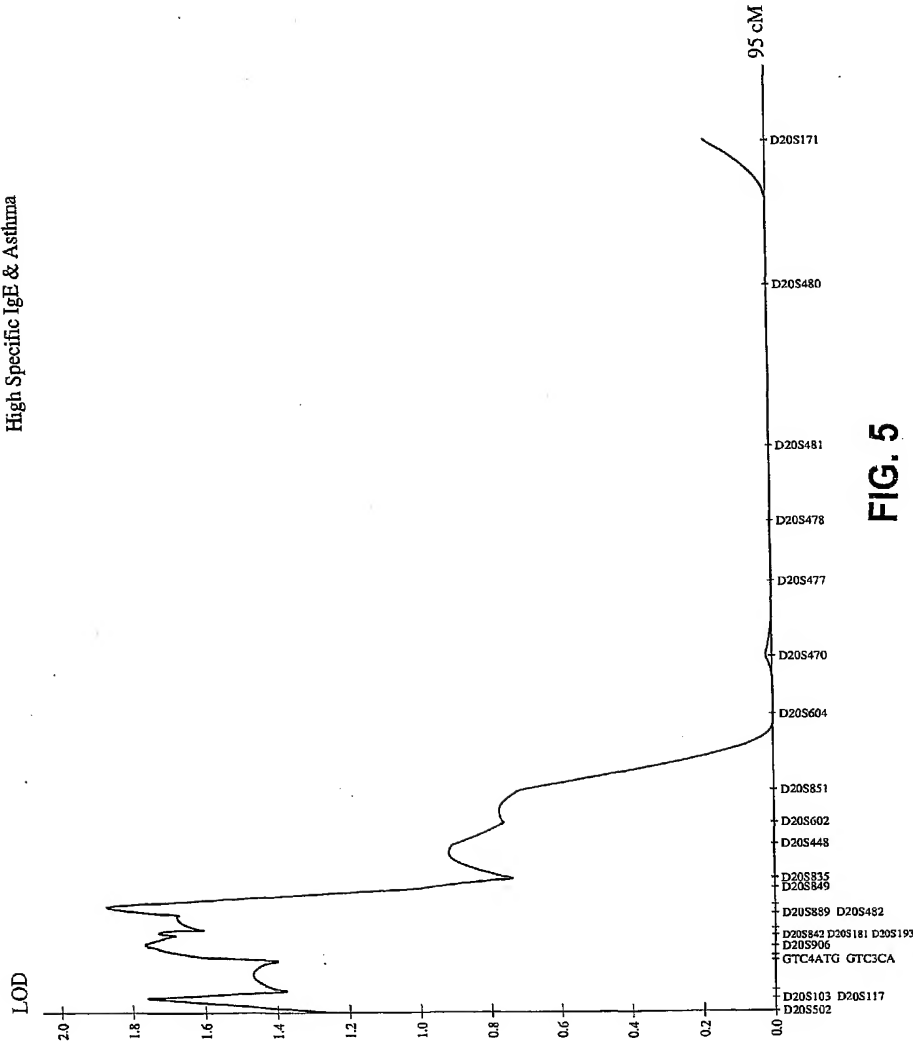


FIG. 5

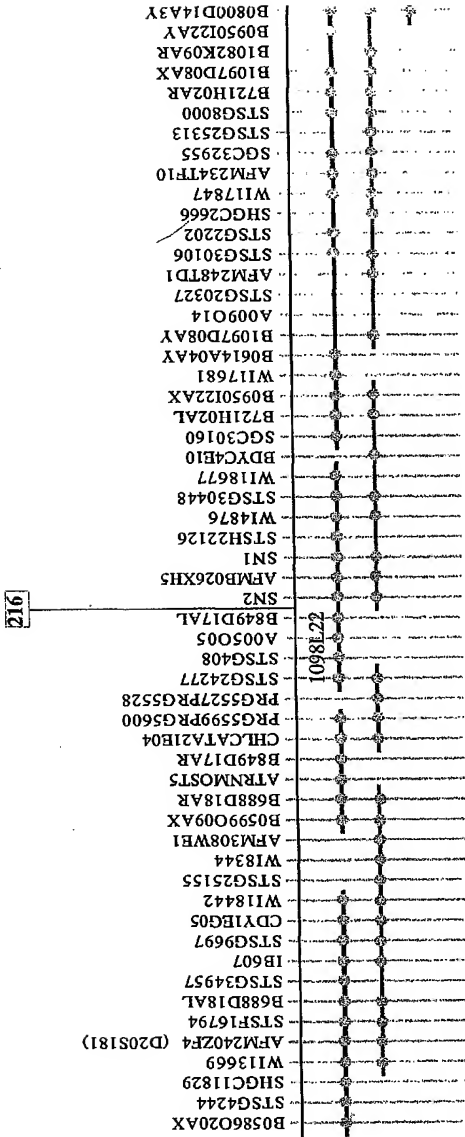


FIG. 6

7/157

B0624E22A2X
B0870L06A2Y
W13772
B768F06AL
W12640
STS025189
RM8ARMB
AFM036YAA3
UT235
STS025337
UT1704
CHLCGATA149B11
CHLCGATA51D03 (D20S482)
B819A20AL
B583H11AR
B715K24AR
B0934N19A2X
B1122K02A2Y
B736H10AR
STS020232
STS020381
W118738
PDG3PDG6
EST498946
W17784
W13651
SGC34960
B727P10AR
B583H11AL
B819A20AR
B715K24AL
B772K01AR
B1122K02A2X
B780N09AL
B736H10AL
B602E18AL
B1033116AY
STS73406
B0686P06AY
B732A06AR
SHGC16916
R79078
SGC35090
W111626
STS08643
STS05017
AFMB3522XD9 D20S895)
STSG42745
STSG10203
B602E18AR

FIG. 6

8/157

>BAC1098L22 sequence

gcttcaataaatttcggcgcgctcaatacgaactcactatagggagaggatccgcggaattcccccatgtgccatgtccagc
agatggttgtaacacagatgagacacacatggtgtgtgcagctctctctcttttttttttttcagacagagctcactctctgtagc
ccaggctggagtgagtgccggcaatttcagctcaactcgccctccagggtcctgattcaagcagattctctctgccc
tcagcctcctgagctagctgggattacaggaaocggcccatatgccagctcaattttcttaggtttagttagagacggggtt
tcaccattgtgacagcggtgggtctgaattcctgactcgtgtgatccgcgcgcctcgccctccgaagtctgtagataca
ggcagtagcacgcgcatccggccatgcagctctctctttctagagttacaacaggagatcgccgacggtctgttagtagccac
aaattctcagacttgcaactcttgaaattacacagcatgacaaaaatgcagctcagcctcagctcctgtctcagttcgtgtg
gtgtgctcactctctgtgccaaattagcttcggtttataactaggaggaattgacttggtgtctcaaatatttttggc
tctctgggctcttttgggtgttcattctatacttagtggattttttttcttggaagaaaagatttcaaaagcaggaatt
tctacttccattatctctacagttgtaaaatttaacctactctgttttctcaactctgtgtactctctcttttgggttaca
aaaacagtggtttcttttagattgtctattttcaaaattatttttaaccagaacatcacaaatatttttctctgcaggtttca
tctctcaattttacttcttaaacctttgcagatttttttttaactctgggtctcagattttgtcctaagtgttggtctcc
agttctagagattttgtttttcttgcagctcttgaccatgacagatgtgacagcacagatgtttacagctgcacagc
caacaccaattgactgccactgggtgcacatgaccatttgctccccaggaaacacagctgtgcagaagggcaaggtcagaggt
ggttcttaagaatttcagaaaaatcccaatttgcataaggtttagctttttatgtagtataaggtataaataatgagggccag
agttacttttcaaatgggtttctattttggatttttatttccctaggttttctcttaggaagagatgttctgtattattt
aagggctcctactagccctgggaattttgtgatatgtgttttagaagaagtgaggactgtgtccctgggcagcggtctgac
taagcactggatccttggtttgagggactgtgatagtgactcacttaccaggtaccaggttttgataaagttagggtttg
agagtgtgactcaggttgtagcagcgcccatatgaaactgtggaaatgacagacttgcgaaaactcctctgttttatactctg
gtttggcataactcctagctctcatatagatgttttaatttaagtataaagtagccaggctttttagttttttg
acagcttacaagactttttttttgtcagctctcatcacagtctataaaatagaaaactctcagacttttgaagtggactt
tgaaaagcacaaggttcaacattccattttacagatgagcaggttgaggcaggtgtgttttaaaagtggggtcgtctcct
tgccagtttaaaagcatgtccttgagcagaacattccattgcaggtgttccccactgcagagctactgcaactctctatct
atctgtctgtctgtgtct
tttgtctctcctaagtattttgtgtctataaattcctgttggtctctcaccataagttttctctgttgatataaaagaa
aagcttttgtctcagctgttctagatctccatttttaggtatgagaattcccccaaggtataccccactgactactgttaa
caagaagaacagctgcaggagctgtgccctggacaagactgcagctgggagcccggaatcaggagtgatgcctcgtc
ccggttaggcttgcaggggtcgtctgtgtgtgtgggtccattctcagcctgcttcccccaacactgtgcagcctaaag
ttgaacctagcagggggaagagctaaattctgtccattctccccacacaggtattatgggctttttgtttttaacta
aaataacgtttcttaagttattgtctctactgcttttgaataaagtgaaacattcttgcctctgtgagaattgagt
acagcttggtttctgtatcactgcagctgtcctgtctctctccccacactttgataacctgggaactgacatgata
tctcagcagctcagctgctgtagcacttttagagacttacttataaaccttgggatttaattgttggcactgctaa
tgggccttctctgagaaggagaggatagatattctgtgcagcagcacttttaggggtgatctccagcctggaacacct
aatatactctgctctgaggttcaggatgatgaactcaggcctgagacagcagcagctgatatatttggaacaggg
tggtccagaaggaggggaacttctgtctgtgcacaaactctgagcttggttcagggaagttgggttggaccaggcccttt
tgaattctctctgggaagtttttgatcagctgcagagtggaacttctctatctcaggtgtgtcctcaatttttaagy
tacaggaagaaagacctgggaattctgcgaaggcatattctgtgcaagtggtggggtggggctgagtgatctttatc
tagtcgactgctacactgtgaggagcaagttctgtctgtcttaaatgatctcttccccagtgaccttttcttact
cagtgactgttactgttaatgaacattgttgatgtctccaaagtacttggtctgttggaagttgctttgttcttcaat
gtcttcagggaagcatctagactctctgcagtcactgtgtgggttcaggatgatcacaggtagcagattacaagct
gtctgtgtatgtctatgccatcagctctgggtgtgtttctcaagaaggacctctcactgtctttggatgctttgtaca
cttgattgtaccttcacactgatgatagaacacagtcagcttttggaactatagattgttaattccttgatctatt
tccattcttgcaagtttttcccccgccctcaaatatgatatacttattgtctgtcctcacttaggattttagttttcta
attgtctctcagaaggaagtgaccagtcaaatatggccaaaactgggtttttctatcaagcactaggaataagtgacc
tcagaataatgtcttttaggattccggggagatcacagctaaacattttaaaattctgtattctgcatttgggtctttaa
aatgtgtcagaggtcccaacatcctaattgaagtacctagaatttaaattagaaagggcttccggtattcagtaatttga
actcotaacagctagttttgtctgtattcttaaaattctttctctttctttccccctaatgagaagaaattctctgtg
catctggctggcatttggttggaacctcatgtttgaaattactactcagggaggaattatgacaaatgcctaattgt
taggagacccaactgccttttggctctcttcaaacccagaagaggtagaattgtccttaagcagctgcgaataatgc
cagtcactcagagcatggtgtgattaaatcctcaaaacttaagtttctgggttctccacttcttaccagggtcactg
cgactgtgactgtggaaggtctgtgatatgtgtaaactcttggcagaagggaggtgctggtgaaacgagctgagga
ctgtgggtttgtgctgcagagctctctctctcttaaaatttttaactatttgtataataagatctagaataattata

FIG. 7

9/157

gggggatgaattgtaacatgtatatgtgtacataatctggtgacatcagtagattatctcatcctgttttacctctgga
 ttctgtcaggaggagaagagaggtcactgataattagctaggttggatttaagccacctgagttccttggagtttaaggtat
 tataatagtgcatagagctgtataat taccactaagaagtgatcatctcagctggatgtggtggcctcaacatagtaaacc
 cagcactttgggaggctgaggtgggtggatcgctgaggtcaggagttcaagaccagcctgaccaatatagttaaacc
 gtctctactaaaaatacaaaaattagccaggcatggtggtgtgcgctgtaatccagctactcaggaggccaggagga
 agaattgcttgaaccaggaggtggagggtggcagtgagctgagattgcccaccgcaactccagcctgggtgacagagga
 tactcgcgtctcaaaaacaaaaaagacagcaaaaaaagaaatgtacatctccttgggtcttctcatagcccgccatctc
 aaaagaagagagcaccttctgtcaagagtttctaagccagaaaggctcaagttctctgcttgcacccagtgctctc
 aggggctcttatagtcataatattccatgatcacattttgtcatttttagtctggatcataaattgtgatctaccgtcag
 taagttagactcatagaacaaagctcttccagcagtttcagctgtggtacagaaacgttttagtggaaatgttcttaccag
 caggaggaggttgaggggcaacacttcttgggtacagcctccttcatgtgaaggtatggaaatgtggcctgggtctcggc
 tgccttggcctctgtgtaccacctatagggccattctgagactggtaggaggtgcccctgtatttagttttctccaatt
 gtccctttttcagtgcaacttagatggggtatggacactcaaacattggtgacatattcttagttgttttacctcaggc
 tactgtgacacatttgggtatttcaataattttgtagctttttcagatttcagaatctccaattggtgactgtctctgt
 tggttctcttccatgtccaaatgtgggtctcttccagcatcccatccttggctggcagctgaccttcccatagttatt
 cactcctcagaaatggatggcaccagcttcttgcctattaccctgggtgtcttcaacaatctctgagtcacaggaag
 ttaaagggttcaaccaggtggccacagataactcttctgtacccttctctcttcttaggttttctaaactcaaacgt
 tcttgggttattctaattctgtctgtggccacgtttatgaacagaaattcagactcttaagttagtttacctcaggc
 acaactgaagtgaattgctcctccacagctatacctctagatgtagccaggttagtagagccagatttttatggaaaaa
 caagaagggacaactagctaaccttaggacagaggtggctggtagaaagctgggaaggaggtgaccctgcaggtgag
 aaggagtgaaattaggtgtcgaagacgggaagggtctctgtgat ttaattagtgcccccatctgtgagatgtagagggag
 atgatttaaggaggtggctctttagtgagctgcaaggttaagtttgcattggttggctgaggtggatgggaggggatttt
 atagttgagctcaggaagggaaccaaagctgcagaccctgcagggccatggctacaatagttaatggatttgaattgtatcgt
 gaaggcaagaataatttgaagggtctttaaaggatatttttaattgtcttctcttcttgggactgtttatttttagg
 tcttttagtaggcatgtgtatttttctctcagaaaaatggaaataggctgggtgtggtggctcctgctgtactccagc
 attttgggagccagtgaggaggatgtcttgagccaggagttcaagactagcctgggcaacacaggggagacctgtct
 ctacaaagggaagttttttaaagaattagccgggcatgatggtggcacaatactgtagtctcggctacttgggaggtgca
 ggcaagagggttctgtgagccagggtttgaggctgtagtggagccatgactgtcctgcactccagcatgggcaaca
 gagcgagaccctgtctcaaaagaaagcaaaaggagggaatacagtatatctttgttttaactacccaaaaataggaa
 tacttaccatttctggctaaactttatatttgatttttaaaacttgttaaaattgcaatgagaaggaaattcagga
 gagcagaagacagactgtcccaggtgctactgtcctatttccctataaaatccagtgccaggatggatgaatggataa
 agcaaatctggtataagttgaatatcccttatctaaatgctttggacgaggagtggttccgaattttagaattatttgcat
 tatacttaaccagttaagcatccctaatctggaaatccaaaatgctctagtgaacattttctcagcatccatcattgttg
 gcaactcaaaaagtttttagatttttgaggtgttttgatttcagatttagggatactcagcctgtgtttgggggtagccatct
 ctctcatatagacatttcagaacttaaatattgtcttgcataaatttctgtgaatttttgatatatttcttctcagct
 acatttttatcctttataaaatggccatatagaagtgtatctatcctaatctacccatggctgagtcagggtataaaga
 ggttttctctgtctgtggggtatacttaacttggtggtttttatctagaagcttgttttggtcaagatgttggttatat
 tcaggccaggcatggtggctcatgctataatctcaacattttgggaggccagggtgggaggaactcacttgagctcaggag
 tttgaaactagccaggggcaacatggcaagactccatctccaaaaat taaattaaaaaaagatcaatctctgtatctatagt
 tgtgtctcttttgccttttagtccaagctcaccttaaccccatgggtcggccttcgggaagatcaatgtgtcctactggtgc
 tgggaagatagtccccatttacaatagtttactacagtggtgcccgtctgagccagtgatgtgtggattctgtggaat
 tttatcagaaccagactcggggactgaaggctgcaacctgcatcaacccatcaatgggtagtgtctgtgaaaggcctg
 gtaggttcacaggtgaaattagggttatcagagtttattgtgagagaacataggaggcatagttcattgtgtgagatg
 tgtgagtagtcatgaaaacagatgaagtattgatttcaagcatgcaaaagaagatataaccagatttcagaagcagaa
 ggaaattcttctgggacccctgaattagttttaaattataagcaaaactaaaaataactcaactactcgaagaaactgatt
 ctaattacaatgagatgtataggtttattgaccagaaaaagatttgagaatttgcctgaaagcaaaatttattgtgtt
 ggcaagaaatgctgtggaagaagaacaaaaaggaaaaataaaaccaaagaaaagattagtaagcaagcaagcctg
 cagggacagtggttcagaaaggtagtgttcaacagggaagaaaaatgatgagagcagttctgtgaagcaggacagaaac
 cagaactacgagagactgcaggagagtagtgagcagacagacactcggagtagctccagcttctcagctctctgggc
 tacatttggctaaactaaaggcaggccagggttagctgtcagcatgaagcaactcaaaataagcctcctcctctg
 cagaggagactcaacagcagtcggtggaaattgtgggttattttgggaggctggacatccatttgcatgtgtccaaatttctg
 gaagtagtaatttggctaggaattacagaggttataggagagaattcatagcattgtagatttctaaattagatttctcaa
 ctctctggtagctgcccagcagctttcagaattctgaaattctgtgaaatttattgtgaacttaagtaagtcgctagaccttctc

FIG. 7

10/157

cctccaaaattgcaaacatagagattttacataattcaagtggaattgtaaaaccocctaaagtctatcctggggtccccaatt
ctaagggggttacagccctacttttatggaaagattcactgtttcctcatctttgtccttaagtctcttggaaagaaaa
tttatttttccacaagtggtgtaatacacttgactaaacagtgacttgatgacttttatgtttctatttctcctcat
taattgatttcattggccacattaatgggtcatcaccocacatttgaaaagtcctggctggggcagtggtccagcctgtaa
tcccagcactttaggaggccaaagcgggtggatcacaaggtcaagatatacagaccatcctggctaaacatagtgaaaccc
catctctactaaaaatacaaaaaattagccagacgtgatggcgcacccctgtagtcccagctactctgggagctaggca
ggagaatcacttgaaacggggggggagggttgagtgagccgaaatttgtgccactgcactctagcctggagacagagt
agactcgtctcaaaaaaaagaaaaagaaaaagaaaaagtccttgacctagcggacatgggtgggagctcctaaagt
tctctcttgggtttcattccagcaaacccacagtgctaaagcagtgccggacaccatgtgacctagggacagcatgtggag
attgcaaccaggcgagctctgagtgcatgtgtgtgacgaacatgaagcagtggtgtggactccaatgcctacgtggcctcc
ttcccttttgggcagtgatggaattggtatacagtagagcaacctgcccgtgaagtgaaggggagccctaggcacttat
gcatgccctctgtataggcaacaactcagccatagggtgtgctgtcagcctctgaacattttagaacaagactggaca
tgacctctgctcaaacctgaccagagactgccatcgagaccttgctgacctaggagaacctcatcacagaatcaggcaca
ttgacagtaaaaataatgagatagatcacagagtacagaataaacttgtccaactcagtgctcatattgctcaatccat
tgaaatctccatctgaatttcccaatttgttaagtaaatgctttccaatagataatctcaaggctccctttgctct
caacatcctgggtagagaggagggaaggtctcctctgttatgtattgggcaaaactcctgggctcttacctcatt
atctcttttaataatcaagacagaataatattttgactcaagccagtgatagatctgttaaaaaaaagttaataca
gtgaattcagatctacctgtgatgtcaatttgcacacttttttttaatagctgaaaattgttcaggctactgtacctg
tgctcatgtctggagcaaccaggctgtggctgggtgactgatccagcaactactggcaaggggaaatgcataggggtt
cctataaaggacagtgagaagtgccttcgcaagccctcagagaaattctatccacagccctgtcctcaattccagcat
tgctcagaggacagcagatacaactggctcttctcattcactgtccaggtgaagtgccttgcatatccaaattcaagtggtt
cactactgatttatgaagaataaaaccttgaaagctacgtttgtgtatagttaactcctgccctcagccctctccttcc
tcctaatgggttggtacagaaggaatagaccagagactgggtccaagcctgacctggagactgtgagagtgtgtgggtggt
tcctaaagaaaccaattctaaagaaattggccttgactcagacttgaagtgaacctcagcaatgtgtctgtgggttctca
gaacagttgggagaggctgggctgggtgcaagactcctcagagattagcactcaagaactctctcaagagctgccattg
acaacagggtgctgttgtgaggactttgtgaaggaaagtcactgtaaacaaagctaaaagggcagagacagctgggaga
aaatcctgcactgcatgtaaacactgtgatcatccagaatagtgaactccacccctcaggggaaaaataaagaagct
aaatgtgaacccaacagaaaggtggttatggagataaagaagccatctcagaaaaggaataagaacaggaaaaatgta
aactaaacaccaggaaaccaatttttgtctactaaactggatcaaaatttccgtgttccctttttccatactaaagatag
ggggacctgcaattctatttttaactgtctaggagtaactttttaacgaaattttaactctctgtttttcatgaata
tcacgaatatatctgggtaaaatgacaaccagagaaatgggagaaaaatattgcaagatatatatctaaagaatccaat
gtccagaaacagtgaaactctaaaactcaacaatagaagacaacccaattagaataatggataaaggattttaatagacat
atgtccagagaagaaaaacaaatggccaaatagcacatgaaaagatactcaatatcatctatccaagaataatgcaa
gtcaaaagcccaatgagataccactttacacccactgagatggctgtaatacaataaacacagggtataacaaggtatggaa
ggatagttagaaaattggaactctcatgcagggtggcgaggctcctcaaaaagttaaatatagagtatcatcacggccacga
gtttactcctagggtacatacccaagaaaaattgaaaacatatgttaccocaaaaatttgtatataaatgcttatgcttata
gcatcattctcatagcatcatgctcaaaagcaacattatctataataaggcaaaagtgaacaaagccaaatcctgttag
ctggcaaatggatgaacaaagtgttgatattcatacagtgtaattgttatttggcaataaaaaggaatgaagtgactgaca
tattttctgacatggaattaccctaaaaaaacatgctatgtaaaagaagccagggtgcaaaagattatgcgttgcatgtgc
catttatgaaatgtccaaagacagaaagtagatgttttagtgggttccctagggtcagctggttaataagaataatag
gcataggtttctcagagtagctgcagcatattttctgacatggattaccctaaaaaaaatgctatgtgaaagaagcca
gggtacaaaagattatgcattgcatgtagccatttatatgaaatgtccaaagggcagaaagtagatgttttagtgggttctca
ggctgtgggagtggaacaaaggaataataggcgagggtttctagatagctgcagtgatctattttctgacgaattacc
ctaaaaaacatgctatgtaaaagaagccagggtacaaaagattatgcatgtgatgcatcttatatgaaatgtccaaa
ggcagaaagtagatgttttagtgggttccctagggtggggatgggaacaaaggaataataggcatggagattttgtagagt
ctgcacattttatgctctcccaacacatcataggggtcatatctttccacatccttgccaactactgttactatctt
tttaatgaaaactgtctagtggtagagaattggcatcctcactgtgtgtttcatttgtattttctgataactaatgaaa
ttgaaatcaactcattagttagcctttgtatctctttgggaataatttagtcaaatctttgtcccaatttttca
gttatgttgtcttttttatattagtggttgtaagagtcttttatgaattctgaagtccttttggaattatattttgcaa
atactttctcccatataatgggtttcttttctcactttcttaatatgccttttgaactcaaaagtttccagttttgtataag
ttccatgtatcagtggttttattttatcccatgtgctttttgtgtatgtatctagaaaactcagtgcccaagtaaccaggga
tcacatagatttgctcctaagtttctttgaaaagtttcatagatttctgtgtgtacattaggtccttgatccatttgag
ttaattttgtgtatgggtgtgaggtgaggtccaaactctctttgtcgtggatgtggctcctgcacattttgttgaaa

FIG. 7

11/157

agatttttttttttttttaaccattgaattttcatggcacatttggtgcagatcagttgactgtgaatctaggaactttt
 ttctaagctctcatttttgcgcagttgaccaatgtttctccttttgccaatattacactgttttgattactgtggcttgc
 tataagtttataaatttgggaaggctaaagctctccactttgttctctattttgcaagactgttacagctattctgggttctc
 ttgcattgttcatataatttttaggatgagcttgccattttctgcaacacaaaagccaactgggttgataaggtttgca
 ttgaatctactgaccaattttgggggtgatttgccctgttgttgttgttgttgttgttgttgttgttgttgttgcagagggtc
 tcaactctgtccaccaggctggagcgcagttgctatcttggtcactgcagcctccgctttccaggctcaagtgatttc
 tccagcctcagcctctcgagtaactgggtactacagggcatgagctaccaacacccagctaaattttatatatttttttagaga
 caggggtttcgccattgttgcacaggctgggtcttgaactcctgagctcaaaagcaatcctcctggcctttgocctccctagttc
 tgggatcacaggtgtgagccaccacgctggggcccttgccattttaacaagggttaactctctagcctgtgaacctgga
 attgtcaatttataccctctttactttcttttggcattgttttatagttttcagtgatcatcattgtctgttccctttt
 ttgggggaaaaataattcagctctttatataattgatatttggtttagtttttttatcatgagtttattgggttgagaa
 agtttccctctatctcagttttttgagggttttgtaatgaatgagtgctagattttgtcaaacactttttctgtgtctac
 ttatgatgatcagaagttttgttctttatataactatgtattacattaaataagaattgttaaacaggattttgcaattcct
 gggataaagccatttagtcaatgatgtatacttttttatcatgatgtatatttttaactctgtgttggtattctttgtct
 agttggggattttgtgtgtacattcatgagggggatataagtttgggttttctcttttggtagattactactcatttttct
 atagcagggtagtactggcctcctagagtaagttggaaagctattcctcctcctatttattacagagagttctgtcaaggaa
 tttagcattaattctcttttagattgtgattgaaatcaccagtgaaaccaaattgggctggccttttctttacaggaaaa
 ttttttaattatataatcaactctgtttgttatagatctattcagattttctttttctggagtgagtcagttttggtaattctg
 tttttctaggaatttgtctattctcagtaatacaattttgttaacatacagttcgttagtatttcttttggtttttttttct
 ttttttttagacagagtcctcactctttccaccagctctggagtgagtgagctgatctctgtctactgtcaacctccgct
 ccagggttcaagtgattctcctgctcagcccccaggtagctgggactacagggcagccaccactcctagctaaattt
 tttattttgttttggtttggcttttttagtagagacaggggtttaccatgtagacaggatgggtctcgatctctgact
 cgtgatctgctgctccttgccctcccaagtgctgggattacaggcgtgagccaccacgcccggccagttctaatgtatttt
 ctgaggggttgcattcttagcaattctgagctctcctgttttaacctgttaacgatttttttctgtctggtattcttct
 tattcttttaccatttttgcaagggtctcaaaatttggagggtttttttttaagagtttgcctctctatatctgtggatgt
 caataaatttaagccttatgaaatgtactaaaattctgtatttttcaattctcctaatcttctgataaacctcctctcag
 gatttgttgcgtgtgctgttctcctgggtggtattcggagattgagagttttctataactgggttaattcctcaagggttaatt
 ttgtattataaactaaattgtgtagatctgtcaaatatttggtcagtggttgggtgtgctgagctgttgggaagttgggtat
 gatgtcctggaaaagctagcaagccagcctcaggacactccagctcctggccactctgactgttgttgactgggtgtgcgc
 atctccctcatgggaagaagggaaggagcaatggacctgggtggctgggtgcaattagcatgttcatgggtgtggcaagtg
 cctccagccttgggtggttagggccacctaaagatagctggcactgtgcttggatttcttctcattctgttctgttgc
 ttgcttttatttttgagatggagtttcaactctgttggccaggctggagtgcaatgggtgtcttcttggctcactgcaaac
 tctgctcctccgggttcaagcgggtctcctcctcagcctcccaagtagctgggattacaggcagctccaccacctctgg
 ctaattttttgtatttttagtagagacaggggtttcactgtgttggccaggctgtatctcgaactcctgacctcagggtgatc
 cactcactctggcctcccaagtgatgggattacaggcatgagccaccacccaggctgagttttctttcaaacactcaa
 atactaacagggtgcataaacagaaacagtaacaggctatggaaaggttggcaagatgggtgaaaaactgcatacactgatc
 aaccaggcgagatctgaagaatcagattttgcttaccagtgcaggttggcactatgtaaatgtttccccatcctatttcc
 cccctagagttctgcagtgacagcctgtgcagctgtcctgagcagctcctcagaggtccaagcttcatatccttctag
 ttgtaaagccctatatcattaggagttcttagtaatccacttgggaaattgttatccgttaagaatgggtatttttcta
 ataatgaaaaatgtaaagcaaaactaaatttagaataaaagggaaatgggttaataactgttaagttatcaaaactcatt
 ctacatttgaagttttgtttataattgtagtaataaagggtggtaaaaatgtttgtgatctgaattgaaagaggtgaatt
 aatatgaacataaattgtgcataattatgttctctgggtatagaaaaataccagaagaaaaatatacaaatgtgaacagtggt
 tctctgagtggttggattgtatgtatgtcttcccatcctctatttttttaagttactgtcatgagatcagcttattttt
 tttagtgaataacacaaaaaataatttttaatatcactatgggctttcaaccttgcctgtgatcattttcaatttttccgtg
 actagctctctctttagataattttttatctatttttcaaattagtttttagtggccagtttttgggtgatttatacaaaacta
 aatttaactacagagatagttgcagataaattctttagaacttaactacatgagttcttccatggtacagaccccttgaa
 actctttttctgcagcttgccaaatgcaacggccacagtaaatgtatcaatcagagcatctgtgagaaggttgagaacct
 gccaccaggcaagcactgcgagacctgcatactctggctctacgggtgatcccccaattggagggaattgtcagcgttaagt
 caaattggtcagggttactcattggcacaattcgggttggaacacagcactctatttgacttgaattctaggagaaaaaa
 gccactttgttttggataccacatttcttataaactatactggtcagaagctcagctgagctggaagaagaagcaacagg
 tttcaggtgtagctggtttaaagcaaacacagccaaaaattgttagagctcaagcctttgaccacattctgcagtggtatagaacaa
 gggcccaaaaacagagaaagtgccaaactccctgttctccttctcctcattgaacaggggacaggtcaatagctcagagg
 gatcagctcagacatgggagctcttggggctcatctcactgagggccaggacaaaagttccagcagtatattgacataaag

FIG. 7

12/157

gttgtgggaaggggcaagggtgttaagggtcaggactgggaaagtcctgggttggagtcagagtcaggaaagtcctcttgaa
ttccccctcctcccccatcaggaggccttggtgctgaagaagacctcagcccaaggccctcagataggaggcctgggaa
taaaagctccagggttaggaatgtgaacatgtgaaggagcgtgacccgtgaggtagacctgagtccttaactgcacctct
ctcagagcccccacacttagggccacacctcacaggtcatgctctgccaggtaaagcctgtggcattgctgtgggtcaggccc
tgaaaatcacacacgggcttttaaccaggagccagactcctcaatactttagtgatttttaagggtttaaatagtcct
cgtgttgggtgggtgtagtggtccagcctgtaatcccagcactttgggaggggccaaggcaggtggatcacttgaggtta
aggagtttgagaccagcctgacccaactgggtgaaaccccatctctactaaaaatagaaaaatagctgggcatggtggcg
ggggcctgtaatccaagccactcaggaggctgaggcatgagaattgcttgaacctgggaggcgagggttgagtcagctga
agatcacgcccactgcaactccagcctgggcaacagagtgagactcagtcctcaaaaaaaagaaaaaaaggt
cctgtgtttccccctttatgtcaaatagaagcctgcaagcaataggacattttatgatgagaaaaaattatgacaagt
gatataatgaactgctttttttattctagtttccagccaaaaaacaataatgagtaactatagtcgaaagactttcaaa
ttctgagcaggaaaaagtaaaacaatagggaaatctctacctctcccagatcccaactccaaactctttggcatggcc
cttcagaaatctgatctgatccagaccagccaggagagaggcagcgttctagctgtgtgctccttactctccctggggtt
cagtttccctcctaatatagggtgggttgttaagatgaaagcagagcacatgacacatggttccacaggatcagtgca
atactcttagtatgttgggcaattcttattacccttctaattgcccattctccactactcctaagtaaaagtctcttgctcaga
ttaaaaggtatttgggtgaaaattgtttccccctctttgtggaatagaacttaacggttaacagagtcctctaa
cagagtggtggagtgtaaaaggcattgtgctgttaccactcaagacacaggtgcttttattatcccccaagatcaagtaa
actgcccagggtctccagctgagatgtggccaaagctaaagaccacgccagtcctgtgctgaaccttaattacttggctgggt
tgtgcttccagatggcatcataccacatcagcagcgtctatagaagcatcacatttctctgtgagatctacagggt
cctgagtcacatttgcttattttccagagtcctaatccacggtgaaaggggtgggtttatagaaatcctgagatggga
ccaagtcctttgatgttttctcttaaaattgtatagtccttaggtaacgatagtagccatgatcttctgaaatgcttactg
tgtttgagacatcatatcatatcatcacaaacacccctatgagataggtactgtgttatcccccatataggacaagg
aagctgaaactcagaaagggttaagtagcatccocagagtgcagagaagagccaagtagaggacgctgaatccaaccagg
cagatggaagctaaaggcctttgctttcactatacatatcacaccttccccagtcctcaacagagccaagtgtagacact
catgatctgacttccagcatccatctgtttaatctgactatctggttttatagttttctttgggtgattccaaagga
tttattttaaagcatacctctcagctgagtgcccttcaatttatccacgcaaaacagggttaaaagtgaaagggt
gttgacaaaaagggtcttcaacttttcaactggagtataatattatccagccttgattacgttggatgataaaagga
gagatgttgagttggccagatgaggagatgggaagagctattttctgttaaaggccaagggaagggtgaagcaagggtgg
catgtctgagggcccttaaaagcctgaccttggcatgggatcctggttctgagccctgcggtgattccaatgtgacca
tcgtggagaaggcactcagccaaagaccagccctccttcagaggacagtcacctccagagcacctcctaggggcagcag
ctatgtctgtgtccccacacccagccactggcatctagagatgcttcacaaatatttattgaatgaatcaaagaatcat
agcagtgaaaaagagagtcattgaaagatcaaaatagtcattgcttcagaaaggcagtggggtaccagcatcactttcc
tctgtcctcttgggttgggttccctttaatgttacctatcatattatcattagctgaaaaatcatgactcttaccactga
atgctcgtactttaaactgtcttctcactgaactcaattcagataaaatgtcgtgttttggaagggtctccaatagtc
agaatttttaagccagttgtgacctgttctcttctctctgcagcatgccaagtgcaatgggcacgcgtctctgtg
caacacaaacacggggcaagtgcttctgcacccaaggggcgtcaagggggacagagtgccagctgtgagtagccatactcc
tggccacacggggaggaccaggaggtgtgcagctgcctgaacccacacctgagagccacccacttccctgtgtcttgg
gctgtgggtctctgagggtacccctgggttgatttgggttgaaatttggccatctatttcagagcaggtcagtcocccaaat
gaggaggtcgtcaggttagcagcaattccttaattggctcttgaaatccacattttgtttaaatgatactgacatttccgt
gggtgtccatttggagtagctcatttaaactcagcaactacttgatttttctcactgcaagatataactcttcccaa
agagtagagatggaagagctgggtgtttgggtgatggcatttatttggcgttgggtgtctgatgttggaatgatcctct
tacctttaaactttccccattcctcagcatcttctctgttgaagcatgtcaggcatcttctggagagggttggaagtc
ccatccatgaggatgcaggtgcagcagcatcagtcatttgaaggttcagacactggccctgcctcactgagcaagggtg
ctcttgaggccaactcactgtccccccatctgctagtagtgaggactagcagtaattgtttttgggtgctcctggtaggttga
aacaggaaaggtatgttccatcaatgagtgagtgactacgggctgtctcactgggaaacgtcagaaatgcatgtatccgcat
actcatctcagaagaaaaacttaaaaagtggttagtgtcaaaacactgagctgataatcattgttaaacatcatatttct
ttaaataacagtagtaataatgcaaccagtaggttagacatcatttctgggtgctactcagcagggtttggctgtctact
tttagttttatctgcataatgtttactagttctatgttggccataaagccaaatgttggaagaaagctcaatccaagg
ttaaatttaacttttaggtccacttcccttaaacagaaaaatgtcctcgagcaaaatataactctcttaccctcagtgta
gctgatgcacaaagacatgagctggcctcaagttaagacttcaagttctcagagtcctgagatcaggtcaggtgctgc
tcatcaccctgggtcaaaagggaacacagcatcctgctgggaaaggccattaccactctccgccttctctctcatgca
tagctgtgtgtttacctaaatgctcatgaagtgaggaggtgggtttctgtgatgtttatagatgtctgtatggggaa
attgtttctgtagtagaacatttttttctacagactgggtgaacactgtttccagggtttaaaaaacagtgattctt

FIG. 7

13/157

ggatttttagtcttctctcacttgtgtttaaagaagagaatttagtttaccaggggaacataggaagaagtgagaaattag
tatttgaaaaatgttttgactgcacttttagaagaatagttagtccacacaaaattgaaaattgttttctattatagtaaa
taggtttattacatttggatttctctcatttgcctttggcacaattgcttgcataattgttctcattatagttgggtgtgtaactg
ctgagcagtcagttatgaagcgttcatgcatgctcttcttagactgctgagctgagttatgggtgaaggatgcaattata
atggctccagactatctgtactttatataaaaggatctgtttgggtttaaaatgaattcagtttctgttttaaatagcag
tataaaatagttctttttccccccagatgtgaggtagaaaaatcgataccaagggaacccctctcagaggaacatgttatt
gtaagtgggttttgcaattcttattcttagaagcaaaagtagttcagtaaaacttcattgtttaaacgggttttgagaattag
aagtgctataagactatagcagccaccaatgaagtggtcccagacttgatattgtttacattctgtttaagtttactacata
taggagcactgttttaagctgtttttaaattgtgttttgggttaaogtttaattgtgtccatagcagatagcaggagagtag
agaggcatgcatctttgtctatccacatttattgttctcttaaaactttactttatttgccattacctagttgggtgtcctc
atatttctgtgttttaggtagtagatcaaaaacagaattcttacaatatgggtgaacttttctcattctgtctggaaactg
atattgtgttaaccaccttgaggtgaaaggtgagctgacaaagtgaccatgtttttagggtaaatgtgttttctcttta
tgggagaatccacatggttagaccagagtacgggaccagaaaaaggagtttaattgttgccatccatttgtaatttat
acctgctgtactgggttgaataacatcccttcaaaattctatgtccacttggaaacctgtgaaatagacttcatatataat
atataaaactcttaaaagaaaaacagagaaaaacatcttgaccttgggtggggaaggattgttagataggattctcaaaag
cataattctataaataaaaagtagactaatgaaactataaaaaatttaaaactttagggaaaaacatcaaaacataa
aaagccacagactagagaaaaactcgcacatcatcatatatttgataaggagtttatatccagaatcatagaggaactc
ttacaactgaataagtagacaattccagttaaaaacaggcaaaagatttgaaaagacatttcatgaaagaagatatcacag
tggccaataagcacattaaaaacgacttaacatcattcgttatttggaaatggagattttaaaacccacagtcagatcacagc
cctaaagtgtattagaatccacgctaatgcatggtcttttagaagacagtggttaattctcatgtctgtcttctgcattcag
tctgttgcagtagacattcttttgggttaaacacatgaaaaaaacctggccttaccagggcatgagttggaaaagggtatagt
gatacccttttaatagcctttcagataatttaggaacttacttgatattatgtcgaactggcaaacctgggtcattttctt
taaagcgggttgatgggttaaccocaaataaaatgatactgcccacagacttgattttcttcttgaggaaacctgggtcagctg
gagtttaaagtgtcttctcgttacttaattataagaaaagcagcctgtatctctaaagacctgttctattttgtgtgtg
tttagtttttaaatatgcaattcttcttccatagatactcttctttagtactcagttcaocttttagtctatccag
gaagtagtgcctattcacacgctcatcaattttgtggctactcctgacgaagtgaagtttttaaagttcttctattttg
ttttgaattttagtaggactcttttcttggctcaattcaggatggagctactgcttcaacagtgctctccagactggagtaga
cgagatgatcttagaggtataggaagaaattgttagactctacgtttatctccttccacataaaaggcaaaagtgtgtg
taataagatttaccaggatcttagacacagactgacattgtattccacgcatcttactctgctgtcagcccatcatggc
ctcatacagaaaggggactccaccatcagaggggcagatagcagagcctgggtttactctctcaagagtgcaccagggctt
aaagacactgctgattgaaatgccagtgtagcagccccaatcagacagcaaggaggggaccccaatcagacagcaaggg
aggggaccccaatcagacagcaaggagggggacactctgctcctagagtgagttcttagcctcattaggaggcaaaacag
caagccttagttgggtccattagaagcttagccagggaataatttaaattgtttaaattgtatgtaaaattgtggatttt
tttatctgctgtcaattacagtagaagcccaacctgacctaaaggtatttaoctagtggtagaaggaaggccacactgoggga
cattttaaaactgaaacatacagaacacgaagatgcaacctgtacagtttcttcaatgaaatataaagtcatgacgataccca
ctcagttatttaaagaaattttgtgtaaacataatggtaaaattatttaggaactccttggagatttcttacttctcatga
acatacacaagccatttaccattacaaaattccattacaataaaattgtgacaaaataacatggaataacataatggcag
taagacaaagtataattgcttggtaacaaatgttattctataaacattgctaaatttaattttaaaagtgaacaaagtct
atagtgtggtatgttttactatacagtaaaatgaaaataggatttggttttaaattcttatttcaaaagataaaatgaaga
aaaattgtactagaaaagatacattctcaaaagttaaattgcttttaagtataaaataaaattactttaaattgaattatt
ttatggaaaaaactatagggtataatatatcatttgagtggtctgattgactggaatttttttcttcaagattttagaata
agatttttaacaaaagtaccatatagaaatgtattcactgacctcaagcaagcgtttgaagctgcaaaactctccagg
atatactctgccttaattgtcattcaaaaaataaataatagtagaattcttggtaataacttcttaaccagggtgggaagca
accatgaaagtatttggacctttctgagtttccagtttacttcaagacagaatacttactagagttatccaaatgcaata
tgttctgtcctctataaaagcacaagcattttaaagtttttgatttcttctgtggaaatataaagtcagcccttgagc
aatacaggtttaaaactgtgaggatccacttacacgtagacttttttcaaccaagcgtggaatgaaatacagttattgga
agatgtgaaacctgagttatacagaggtggacttttcatatgaagggttctatgggcagactgoggactgaggtattgtg
tggattttgggcatgctcaggggtgctcggaaacagtagaccacagttaccaaaggtagactgtaattgttagttgtgtgc
tgccagcaaatataaaatacaaaataagtaaacacttgaagctgtccttcaagatgaaggtgaggtcttatcagtaaatga
gaaggtaaaatgctttgtgagcaaacctctgtgtttatgatcacattttaaaatagctgtttggaaatgttccattgt
tgtgtattttgtgtctattaaaatgatcaaaacacacaccttaaaactcttatttcaacctcctcaagatcttttaaaatga
ataattttagtcagtcagtgagtgcatctgtaaaagataaaaaataaacattgtagtttgcaaaatgaattgtttgacct
cagttaaagtagtactgcaatttcaagtttgactaattattaatgttataatttacctaacatcaccaacagtagaac

FIG. 7

14/157

ttctccactggcttaacagatctctcagcaatatctttattagtcattaaagccaaggatcaaaataaataagttaga
 ttagccctgtgaactgctatctctaaagtggtagacaagggtttcaaaactaagaagccatctcaatcatatttctct
 tgaataaataatttttagtaagagaaaaatattttcaaatcatgaataataaaaaattatttttaaataggtacattt
 cagctctatatgctcttttttaaaggctgtttccaggttttgagaaagttttacactgtataataatctatgagtttagaatt
 atatgggttctagttataaataaataataattttggggcacttgatcaagtggttttattggtagggagtgaaactcagacatt
 tgaanaacagctgccagagaaatctctaggagggtgattccccagcacaggtcaggagatgcaaaactcactgtgctgata
 ggggtctttttaaaaggcactttctcatgaggttaggcatttgataaagcatataattatgtatacttttaatgggagggga
 ataaaaattttattggcatatattgctgttatagggaataattgggttagaataatttgcttaactctctgtgagtttctcta
 aagacatcataaaatggtttgtgattaaagaaacttagaggagtaaaactttattgctttattttaaaaatctagaataattgtt
 ttaatttaattttctataaatttgtaacgctcatcagcaaaaacagggttttggaatggttcatcaatgctccaagaattt
 caacctcaacatcacctgggctggcaggtttctcaggtaaaagacatacctagagaagaccctgcaaatgaaggtgtggttag
 attaagaataatgataataggaattgagaagcgagctcaggagacagattgggttgaacccacaccttgccactactag
 ctatgagaccttgagcaagtatctaaatccccctctaaaccttagcattattttatctatctgtaaagtgaggataattga
 tactacctcttagaattggtgtatagattaaattaggttatacctaccagagcttgctgtgggtgtagggctcagtggtg
 ggttactactagagcccaacccacacctttctgttctgtgctgtggccactggcactaccatttgctctacatccgtgct
 tcggaaagtgaataattcaaatgatttgctttcaataaataaagaaaaacattttaataaaaaatgagatttttagtaggtacagaga
 aatgtaacttggaatatacacaagctctaaaagcacagctctgtgctgtgcttactgtgattcactgaagatctactct
 gtatagaaaaatcaaaagaaataaaggatgaaggcccaagtgctgtggctcatgaccttaactcaacacttgggagagctg
 aggcaggaggttgcttgagctcaggagttttaagaccagcctgggcaacatagttagacctgtctacaaaaaataaaaaa
 aaaaaaacccagggttgatgactcatacctgtaattgcaactactcaggaggctgagatgagaggatcacttgagccag
 ggggtgggagattgactgagccatgatcgacaccttgtaactccagtggtgggcaaacagagcaagatccatctcttaaa
 aaaaaaataaaaaaagaaaggaataactactacttaacttgataactcaacagtagataatgggtttgaaatttgctc
 catataatttacttgacccgtgctgttaggcaagttacttaagggtttctgagccaggtttctactgtataaagtaggata
 gtaaaaaacacgctcctggcaggcggtgatagctcatgctgttaactccagcactgtgggaagcccaaggtgggaagatcac
 ctgaggtcaagagtttgtagaccagcctggcccaactggtgaaacccctgtctctactaaaaatacaaaaaacagccagc
 gtggaggccagctgacctgtaactccagctcactcaggaggtcaggtaggagaatcgctgaaacctgaaagctgaaggtgtg
 tctgagccgagatcacgcaactgcactccagcttggtgacagagtgaaactccatctcaaaaaaataaaaaaataaaaaaac
 acctccccagtagagtgatgtgagaataataaattgagataataaataaagtagtactcaatagtagtctgaaattggttaatt
 ggtaactatattttatcatatttactattacaactactgggttttttaaaaaataaaaaacaaagcaatgagatgtatgca
 aaataagaataatgctctgtgcacgccaactacgtttatcatcttaaaacattgtgtagaatttgagaaagttcagaaaa
 ctctcaatgaggaggagcttttaagaaaaagttcgaattatcagagattttggagaagggcaacatctccaggcatgtga
 agagattgcaatgagccggcggtggtctatgctgttaactccatagcactttggggaggctggggcagggtggattacctgag
 gtcaggagttcaagaccagcctgacccaactgggtgaaatccgctccctactaaaaatacaaaactagctgggtgtggtgg
 ggcgtgctgttaactccagctacttgggaggctgaggcaggagaattgcttgaacctgggaggcgagggttgagtgacc
 caagattgcatcattgcaactccagcctgggcaacaaagcgaaattccatctcaaaaaaataaaaaaataaaaaaagatt
 tggcatgagtgctcaatgaagacgtgataatgtgggtctagtcacagggtctaaactcagacatggaaaaaagtcattt
 tcatataactctttatcgccacttgaaattctctggctcaagggaagattggaacatgaaggactctctgggaataggatgga
 gttataccagattagggggacttaataactctggtgagctgggtggtagaaggaggagctgagtgacccctgaaacctctcc
 tccctgtcagtggttgtaggcagtgagcgtgtacatcagcattactggcatgggagctctggcgcattgccaaggagggtgt
 aaaggggaaatgcaagggaattgaagtggttggggcaaaagtgaatgccagtgcttgttaataggaattgagtggtatctgt
 tattttctatgatcatgtgtgtcacctgtttgggggtggggcaagggtggaaggagattacatggattctctggttaaaacca
 tttttcttcttttttttttttttttgagacggaaatttcgctcttatgctcaggctggagtgcaatggcgcaactctg
 gctcatgcaacctctgctctccaggtttcaagcagttctcctcagctgctgagtgagctgggattcagcagcatctgag
 ccacacgcctctggcttaatttgatttttttagtagagacgggggtttctccatgcttgctcaggctggctctcaactccgaact
 cagggtgataccacccgctcgggtctcccaaaagtgctgggattacagatgtgagaccgcacccggccagctaaaacctgt
 ttggttaggggcataggcttgatatacagcctgcccagctttaaatcttatctccattctctgttggtgctgtgacttga
 gggagttacttaattctcttgaaacctcaattttccagctctataatgaagataataataggcgagctgtgaggatataaa
 tgaattgagctcttaaggctattctgggaacacaggaaatttgataaataatgctattgtctcttatttataatgaggcc
 agattctgtctctccctcagcccccacaaatgtctctctctctcatgtttttcttttaacagctgggaaccaggctgt
 gagaagagatgctgtgtgtttcaaaaaacaaacttaaggagtagaagaatgattttctcttaattgagagttgtatttctgc
 aaccacccaaatatacatttctttttttatgtcagtaatttccactggcccatcaaaatcagggtgaagaaggagctttgt
 gtctcatactgcaaaaggttggtgaaattctttagtaagactaaattactaatttggagcttggtgttaattgagatgtg
 caatgtggctttgctcttgtaacgtgtatggcaggagggtgctgagcacatgcatgctgcacagaagattggagtgggga

FIG. 7

15/157

tggactgtatcactcatgaaagacatttgcataaagcactgttgaaagcaagttggcatgtaacagatttggctattataa
acgtattcacttcttcagtgagcatttggccattgtgaaagcctgtagggtacacaaagaacttcaattctagagtaaggt
ggatgtaagtgaacaaacccacataatacaactgaaagccagagttggaagtgacatgagatgatcccgaaatgct
atcaaaagttaaaggaggacaaatggggagagctattgttgaagaacatcagccttctagtgcagacagaggtggattgtctc
ctgggtcctatacaaatcatacagcctttccaagtctccagtctctgtgcatgtgacctgaagggtagttgttaeggggga
tgtgacgtgctgcagtgcttctgttagctgctcctttctctgttcccagggggccagtgtaactccctctgtgcogaga
cccatggcccattttaacttttatactcatgtcccttggggccttctcactcaactcttctgcttcttaccttcttcat
ttagggtgaatgtggaggttagggataggtgggcttccaaggactgggtcacctttaaccatggaagcatgggtcactggac
ggaggctgtgtgctgttggccaatgttcagaagcataatcaacctcagaagcaagtcaccacaaacatataaaaaagttc
aacatcactgactcattagagaaatgcaaatcagaaccacagtgagataccatctcacaccagtcagaatgtctgttatta
aaaagtcaaaaaagaaacagatgctggcaaggctgtggagaaaacaggaaacgctttacactgttgggtggagtgtaaat
gttcaaccattgttggagacagctctggcaattctcacaagacctagaggcagaaataccatttgaccacagccatcccat
actgggtgtatatacaaaaagaatataaatcattctgttaacaaagatacatgcacatgtatgttcaattgcagcactattca
cagtagcaaaagacatagaatcaacctaaatgccactcagtgatagactggataaaagaaatgtgggtacatatacaccatg
gaatactatgcagccataaaaaaggaatgagatcatgtccttgcagggacatggtggagctggagggccattctcctcag
caaaactaatgcaggaaacggaaaaccaggtaccacatgttctcactcgtaagttggggcggagaaacacatggacacatgggt
ggggaaacaacacacactggggcctattggagagtgatggggagggaaggagagcatcaggaaagaatagctagtgagtgctg
ggcttaataacctaggtgatgagatgatgtgtgttagcaaacactctgggcacacgttacctatgtaaacaaacctatacat
ctgcacatgtacccctgaacttaaaaaatcaacaatacaaaaaaaagcaagttatcaactcatcaattaggatgcttggg
actgtgactaaagaacagttgatcttccattctggaaaatattggaggacaaaaaactatgtttagtttttctcaccaccc
tctctccttttctgtctcctcctggctcaggttcttctcctcctcctcctcctcctcctcctcctcctcctcctcctcct
gtgtccctcctcagccagctctgtgaagcatcagccacgcttcttattctcctgttctcctctgtctagtcaacacatctgtc
acaatggtctggcctcttccctcaccacatcctgaaactgtcctgtcatagtgggcccccagtgatggaaattttcagttc
cctttttatctcattgtatggcttgaatcatcatttctgttctcctggatttcttctcctcctcctcctcctcctcctcct
ctcttttagggaaatcctcctcctcctccttaaacactgggctgtctaggattcagacactggctccttctcctcctcctcct
tctgttctcct
tgctgctggacagccctacctgggtagtcaggatccccgcgttcgggtgcatctcctcctcctcctcctcctcctcctcct
cttccatgcttctcctcctcctgtaaacagccaccacatcaccctgtgtgaacagtagggcctgatgatcatttccattcc
tgggataaacccaaacttgagcatgagctatttcttactttacacgttggtggattcagtttgctaatatttaggtt
agaattttgtatctatttcatgagtaagaatgtctgatatttcccttctcctcctcctcctcctcctcctcctcctcct
aaactaaccttttaataattataatttgaggaggttttccactccttctcctcctcctcctcctcctcctcctcctcct
cctcaagtaacttggtagctctcctcctgtaaaaatatttgatcctagtgcttatttggaggagatacttttagctgtgta
ttcagctccttttaatagttataggtattttgaatttctccttattttagtccacttttaatattttctcctcctcctcct
atatcatggttttaaatgtatttttaattttactttttaaattgatacataatttgatcgtatttatagggtaacatgagat
attttgatcatgcatcaaatgtataatgatacaataagagtaatttaggatccatcactcaaacatttcatcttct
ttgtgttgagtaactttcacatcttctagctattttgaaataatgaataagttattgtcaactatagtcataatgactat
tgtgccattgaacacttgaacttattccttgtaactctattttatataccattaacctctatctccccagccccagca
gctggtaaccacacattctgcactctacctccatgagatcagtttttttagctccacacatcagtgtagaanaacatctat
cttctctgtgcctggctgtgttcaacttaacataatgaacctcagttccattcatgttgcgtgaaaggacaggttccatt
cttttttgtaactgaataatatttcatgtataataatacaacattttatttgcattcatttgttgaaggacagctag
gtgatctcagctctcactctgttgaatgtgctgtgataaacatttagagactgcagtatctctcagtatactgatt
ttctttcttttggtatataccagcagcggtatgctggatcatgtgggtgatctattatgagcagctctcctcactgtc
tccatagtcctgtactaaaataatttacaattccacacagcagtgaaactagatcgtcttctcctcctcctcctcctcct
tctgttattttgtcttttaataatagccatttcaactgggatgagatgatttctcattatgggttttgatttgcatttgc
ctgatgggtactgatgttgagatttttcatatgcctgttggccatttgatgtctctcttggacaaatctcattcag
atcatttgccatttttaaatcaagtttttctcctatgagttgttgaattgctgggtatattctgggtataaatccct
gttggatggatagtttgaacatatttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct
cttttggcttgatataatccccatttttggcttttgggtgagatttgcacatctaacacacatcttctcctcctcctcct
ggccagaccatgtcctgaaggtatttcccaatttttctccttagtagttttaggttttaggtcttacttaagcctt
taactcatttaaatttgatttttgatgtgggtgagagagaggggctagtttatttctgcacatggatccagttt
tccagacacatttatgaagaacctgctccttccccactgaaatattcttgactccatttatgaaacagcttggcgt
gaatattggtgatttatctcaggttcttatttttggteactgggtctgtgatctgttttctggtggtatcagctgt
gtgggtactatagctttgtagcatatttgaagtaaggtgatgatgctcagttttgtcttcttgcagaaatgc

FIG. 7

16/157

tttggttggtgagtagacagcagctcgtgcctataatcccagcactttgggaggccogaggctgggtggatcacctgaggtca
ggagttcgagagcagcctggccaaacttggtcaaaacccctgtctctaataaaaaatacaaaaattagccatgtgcagtggtgg
gtgctgttaatcccgactactttgggaggtcggggcaggagaaatctctgaaccgggaggaggaggctgcagtgagccaa
gattacgcccactacattctaccctgggcaaacaaagcgagactctgtctcaaaaaaaagaaaaaaagaaaa
tgctttggctatttgaggtctctgtgggtccatcaaaatttaggattgtttttctatttctgtgaagaatgtcattg
gtgtagagatatacttgatctgtgaggtagcttttggtagtatgggttttttcacaataattaattatctcagccacaaa
tgtgggtctctcttcaattttttgtgacctcttcaatttctacatcagtggttttatagttttccctgtgaagagggct
ttcacctccttggttagattttatctacggttggttttggaggttttaaaaaatctatcgagaatgggattgtcttct
tgatttctcttttgctagtctgtgctcctataagaaatgattctgatttttggtgtgtgattttgtatcctgcacact
ttactgagtttgtagttcttagtttggtggagcttttaggggtttctgtatataagattatcaactgcaaataggga
cactttgacttctctcttccagtttggtgacctttatttcttcttttgctagtgtctctggtcaaatatgttgataa
ctgtgtgagctgtcctctaattgctcgagcgtgtgtgtgtgtgtgtataataaaatgttattcagcattaaaaaggaaatctga
tcacatgcgcaacacagtggaaccttgaacacatgctcaagtgaataagctagttgcaagagggaacaaatgtgaatgat
ccacttaacatgaatatctagagtagtacacagatctcttagagacagaaagtggttgagggtaccagcagctagggg
gagcgggaaatagggaattactgcttcatgggtatttaaagagtttccctttgggggtgatgaaaaagccttggagtaggt
agtgatattgggttgcaaaaaattgtgaatgtaattatgcccactgatgtgtacacttaaaaatgggttaaggtgattttat
ggtatattctactacaataaaaaagctcttaaaaatctcagcaagatgtttttgagatagacaagcttattctgaa
atttatatatgaaggggcaaaagtctagaatagctataaacaatttagaaaaagagaataaaaatgggaagaatcaacctac
ccaatgttaaggcttactatttaggtacagtaatacaagacattgtttcaggtagagggttagacacacagatcaatggga
taggtgtatctataggttagatccagaaaatgaacacacataaaatattgtccaactgattaggatttgattttgtttgtggt
gttttttttgagacaggggtctcactatgatgtccaggctgggtcttgactctggggctcaagttaacctctcacctcagct
ctcgaataaactgggattacaggttgcaacacggccagctagttccagctgaatttttttttttttttttttttttttttt
aggggtctcactttgtcgccaggccaaaagtgtagttggtgcagtcacagcacactgcaagctcaactcctcctggctcaag
gatcctcctgcctcagctcccaagtgtgtgggagctacaggtgcataccacactcagctggctcaactcctgggctcaag
tt
tgatcctcccgcttggtctcccaagtgtctaagattacagaca tagagccgttgcccccagccccagctgattttgac
aaaggccctgcaaaaggcaattcaatgaagaaggatagccttttaacgaatagtgtcagagcagttggacacacataggc
aactaaagggaatctctaactctcacatcttatataagaattaaacgcaatattgatcacagatctaaatgtaaaaacataaa
actataaaacttttagaaaaaataatagggaataatcttgaatcccaaacgaaacaaaagctttttgctctgtgagaaaacctgt
aaaaaacctgatttataaaaggaaaacttgataaactggacacataaaaccaaagtttttgcctgtgtgagaaaacctgt
tatgaagtgaaaaggcagctatggactgggagaaaattttgcaaaccaattttcagttagaataatgggcaaaagacatgaatagacattta
tataataactctcaaaactcaacagttaaacacaaagcaaaatcagttagaataatgggcaaaagacatgaatagacattta
ccaaagaggatatacatatggcaataatgacataaaagatgatcggaatcaataacccattagagaggccaggtgtggt
tcacactcgaatccagcactttggggaggccaaggcagcagatcacacttgaggtcagggattgtgctactgcaactca
ccaacatggcaaaaccccgctctctactaaaaatacaaaaattagccaggatgtgtggacatgacctatgtcccagctac
tcaggatactgaggcaggagaatcgcttgaaacctgggaggcagaggttacagtgaggtgagattgtgctactgcaactca
gctggggcaacgggaatgagactcgctctcaaaaataaaaaataaccattagagaaatgcaaatgaaagcacaactga
gataacactacacacttatcaaaatgactaaaaattaaaaatagtaataacaccaaattgtggagagagttggggaactga
ggatcactacatcattgtgtgtgggaataataagatggtacagtcactctggaaaactttgttctcttaaaaactaatg
ttatacttatcatataacccagcaattgcatctcaggcatttatcctagagaaatgaaaaacaggtgttctcacaactct
ctgtacacccgtgtttatgacactgaaagcaatatttaacagctttatttatactagcccaaaaactgcaaaacagc
aatatctcaacaaatgaatgcttaataaatgatgctacatacacaccatggcctaataccagcaacaaaaagaaatg
aactgttgatatacagaaacacaaacactttgataatcccaagcaaattttgctgtatgaaaaagaaactctagaat
ttatacttgcatgtattgcatttatataaaacttgaaaatgaaaattttagatcttatatacctatgcacatattgata
agtatcctgaagcttcaaatttctgactctcttaggactctgatctctagacccctgttttaaaggagcttactgtta
ggtcaggcctaccacacaggatagggaagatcacagcggtgccatcaagaagggtgggaactcttgggggcattctta
gaattctcctcgtgtgaggtgcatttggaatagagcttgcaagggtttcggtgatttctccgatccgaaaaagggttag
ctctaggtcccgctcagttatattttaggttggaacacatctcctgacactataagatatagatatattgttag
gttattttcagttaaaaacttgatttaggtttcagtaataaaaaattttattttactcattttaaggtgagattgttt
tgtactgacatacctctccatattctcaactcaattgaacatcgtgctttgtgtctatagctgcaaaatgatttaact
ttttagattcttagtgaaaaactctgacttatgtcgtgcagacactttccagctctgtgcctataagcaggtctttgtgctg

FIG. 7

17/157

gagatggccactcttgttgccaatgccagagtttttttctcccaggtccctattgctccaccctctgaaagtcactgaa
tggttttagagctttcattctagagggagctctcccacacttactcttaagaagtaataattttaagtttcataatca
aggagatgttgattttctctcctgggtttctgggtgctccttatctttctggatcactaacttaaatagatccacttat
gaatgctgcattcaggagcactctgtgaaaaataacgatggcatgcccaccatttagtttgccaaagcttccctgaaaaacg
caggttaagttctattgcaaacgtgtcttttggctgtttggtaacaaatgaotgcactattcactgtgactaggcccggtgc
acogtgtgttctcagtcctcgcatgagagcctgggtgacactgggtctctctgctttggggagccgtatctggctgtgoc
tatgtggagggtcagctcgtgtctctcctgggttaaatacatcatgttctcaacagggtgcagttctcttattgctgt
tttttcccagagacttagcttttctagggtgtgtagcatggctcttaatacacacatttagcaactgtgcaaatcagt
aaaaatggagagacttatttaagggaagacagtagggaatgtctaaatagtgacgttagggccattcagttaaattcag
tatggctttcgatgcaggcaaaagccggttttgcccttgattctcattttttaaattagattttaaaaattgagattctcct
tttttaaatgagattttaaaaattgagattctccttttttaaatgagattttaaaaattgagatttctccttttttaaatg
agatttttaaaaattgagattctccttttttaaatgagatttttaaaaattgagatttctccttttttaaatgagattttaa
aatgagatttctcatttttttaaagcaaaagtaactgaggaaatctaattctctacttggaaatttggtgtaagcattgggg
cttttaaaaattggtaataactaaactcaagttcccttcaactctgaaccaatcccaggattcttttccataattagggaacagat
caaaacctaataactcgtgagggaatttgaactgcactataaaaaactcttgtagatgcagctaaaaaacaaaacaaaccca
cagactgaaaagtcctgtctcgggaggggtggcggtgtgttggggccgggtgtgtgctgttttttgattgttctgtctcctg
tgggtgagcacccttccctcctccctgtctctgaaagctatcctagttgttctcccagtaagccggggtcagccgg
cttcagctatgtccatagtttgggggctggggcggttaggtatcagaatagaagatggttggaacagaagtggtgtgaggt
gggaagggtggagtggtgaggtctgggtcgttcccgactctgcaacagaggttgttactgtaagtggagctaataataga
agctgcgtgctgggtgttctcatctactgttttatgtaaggagccagtggtcacaaagtaagggaacacagaaatgggtat
tgagaactgggcatctattagatttagagttctgttttatgggtgtcaacaggaatcgctactgagcaagggtatgtgt
gaaacttttgttcaactggatgtgtcttttttctataacagcttattagagacactttacataccatacagttcgtc
tatttaaggtacacaattcagtggttttcaggtatctacagaagcgtgcagccaccgtcactacagccaatttttagaa
cattttcaccaactccagaagaagccctcagtcagtcagtcctcactccccctctctccggcaacccccagttctgtatttc
ctgtctctgtggatttgccctattctgggacattttatataaattgagtcacatgatagctctcacttagcatactgtttt
caaggttccaccctgttttagtatgtttcagtagctattctctgtcatagtgtaataatctctcttatatgatataccac
gttttgtttatccattttctatcgatggatgtctgggtgttttcaactcttggctattatgaatagtgtagatataac
ttttaatgtttatatttattgtataataacagtgcatctcgtacccctcctgggttaattgtctgtatcttaagagtttt
aagaattataggttgtgtgctgtaattctagcacttagggagggccgaggccgggcagatcgctttagctcaggagtttgaga
ccagggtgggcaacatagtaaaacccctgtctctactaaagtacaaaaaatttagctggggtgggtgaagtgttctgtagt
cccagctacttgcaaggctgaggcacaagaatcacttgaacccaggagggcagatgttgagcgagccaagattgacccac
tgtactgcagcctgagcgacagagcaagactcctgtctccaaaaaataaaaaaattttctttatgatgaaccccaag
gctgttaacaaacctaataatagtttgaacacaagaagacacatgcttttaagtacagtttaattctctttctttctc
ttatttcattttttttttttggagacagagttctcactctgtcgccagggtggagtagagtggtgcaactcttggtctg
ctgcaacatctgctcctgggttcaagtgtattctcgtgctcagccttgcagtagctgggaactacaggaacacttaacc
acgcccagctaatttttgtatttttagtagagacaggttttaaccatgttgcgcccaactctcagcaaaactatcgcaa
ggacaaaaactgaacacogctgttctcatttcataggtgcaaaactgaacaaatgaggacgcatggacacaggaaggga
catcacacaacagggctgtgtgtgggtggggggaagggaatagcatccaggggatatacctaattgttaaatgacg
agtttaatgggtgcagcacccaacatggcatatgtatacatatgtaacaaacccgcatgtttgtgcaatgtacccataaa
cttaagatataattaaaaaataaaagaaaaatgtgtattctcttgtctcgaaagtcctcctgttttcaagcttgttctg
ttgttctctcttctgagctctttaaagttatgtttccctgggtttctatctggactcttttccctgggcaactctcatc
cacacactctgggaacatttgcccttgggtgacaggttaaatcttctgtctagcagactctctcctgtgactccagag
ccaggaacccactagaactccttaaaattctatcagccagcagactctctggccagcagcagcttctggccagcagcagc
aggcaccctcgttagaagcttggtgttggaggtgtgcccattgctcctgggacgatgagagggggggagtgccctgt
gtcactcagtcgaagttaagagagatcacaccagagctgattcctaaggagcatggggggggggggaagggtgttctgct
ggtggggaaggaaatgacccgctcttcaggaaactaccaggagtttagcagagctggtgacagtgctatgagcaagaag
gtggagagctggggaggggccattatgaagagactggagattgcaaggctgtaacttgagtcctgattgactgggaattg
ttgagatgttttaggcgaaggtcacccagagcatcagttttagaaagacacagcagcttgggagtgtagttgtagg
gaaggagccagggcctaagtctggggaggaacttaggaggtccaggaggtccagggtggaattctgtgttagaaa
cccgggaagggggacactctgtagaacacacatccatagcaactctcttctcctctgcaactccacagcagagtgctc
gtttccactatacgaagatttcttaattggcaaggatcgcattttataccctcagtgctcctagtgctcatgctaatcag
tggagacactccctaattgtagcagatgagatagttgttagtataagggtcttaacgggttaggtacagaactttt
ctctttttacaacttttataaactttataaaggagccgacacottttctgatattgtcagaacatctgctgacagtg

FIG. 7

18/157

FIG. 7

19/157

caaacctcgtgggctcactgagcttccgagttcagagtttttcagaattgatgaacacatccagaagggcccatggggcaga
 tacaccttagacatattgatgaagtgagggtgggaattgggttgggaagtcgtaggacgtgaattctagggtttcttggga
 ggtgaggagtaggtgtctactgttttgaaacacagagaagaatgacattatgagttttattataaaataagaatttctct
 ttgtattaaaatagatttttatataggaaatttccattaccttcagtatcagaaagggccccaatctatttccacggtt
 ccataataaattgtgaagaaagatttgattggggtattgtgctttaaaaatgaaccacaaaagtaggggaaaactctta
 atttttagagtttaacaacatttatagttacaatcatgaagcaggtatgggtgctgcttcatgagccaagttcataaaagaaa
 aaaaaatcagatcttactgagttctgtgagacagatcattattttcttataacatatctatttattcagcaaatggtt
 attaatgtccaattaatgtatcagtggtgcttttaggcactggagatatatgggtcagtgaggatgaaggggaagttctgtct
 ttcaatcctcagagatcttatttccataaagtaagtgacaaaatcggggtcagtaataaataagagcaagataaagagat
 ggggggtggtgcttcagtttggtgacatgggtgcccagaaatctgtagctccatcatggagttcttacttgagacttcaca
 gcaggtatgtaaaagcaacagccaggaaactccttagtttcccaagtcctccttaactctggtttagagaattatgaa
 acacttctctatggcttcatccttttcccttctcagttgataaactgaagggtcaaatctccgactctcctttgccc
 tgggtttgtgagcaggttctccaacagcgtctccactaaaatctgtcttttacaagatggattttacagttggttggc
 attgtatgcagctgtgtagtagggacagccaattgtattgttgaatacaagtttatgcctttacatttgagtaatttt
 taattacctcgtgtaatttttaacagcactctccttcatcttcagaagttggtcagtttgaaacataagtttgtcatctgt
 tcagttcaaaaggttgaccaacaaagactgacccagtcagtgggcggggtattcagcaggggttagattttctaaagttcagtg
 ctctcataaacattttcagcttatgacctttgctcctcacacaattctgaacttaggagacttagatgataaagattc
 aactaagaggtatccctttaaagttctgggtaccagccatcccaaaagctgttttatatgaggttctttatgattataatt
 tggcagtgaaacatggagttggtcctgcttcagcagaatcagggtcaataaaccattagcatccagttggtgaaagcggg
 gggtgctcctgaggagttgtgtgtggcacagggtgcatgaaacacagataagtcctcagacagtgcccttagcatccagtg
 gtttttggcatctagcagggagtaaaagattctcatagattctttaaagtgtaaacacccagggacacacttaacgttttaa
 aagattgaaatccagacaggggtgggtggactctaatctgttctctctctcttattgtgtataaacttggtattctt
 cattccccaaaatgacctgttttatataattctccagctcttatctaaaacttttatgtgaggaccatggctcattctcg
 cagcttctctctcagagagctaggactgggtattgtgataaataagggaaaagaaagaaaggggtttctgtttat
 tgtatgaaattgatgtaactttgtatttctgtattataaatggtagccagggaggacacaaaacttcaaatccaaagc
 cccataatatattatggcctgaatagttgtgaattttattttatggaaatgtcacaagaaagaaatgaaggtttaa
 agcaaacagctggaagcttattttctggccctgtgagctctggctctaaaaacagtgcaaggggttgggggttgggtgggac
 taggaaagcagatgctcgtcatttaccttctgtgtttgtcatgagaattgtctgtactgggattgtcaocttgggtcagt
 aattgtttcgtttatagtaaccattgtctctccaaagaggggtgagttcctaaaaaaagcagggcagccacacacat
 ctggcttagtagcagctaatattcgggtgatagtaaggacacaaattctcgtattttaatcaatttggagggtatgcattttt
 ogtgggctaatatccctgaatcatgatgcattcttaggtcagtgacttaaaagcatttatgtcatagttaatttgacag
 tgtttttccctaaattatacacagaataatgccttttccagccaattgtatttttagagatttgatgaaattatgttgcatttt
 atattcttgggtatgaagactatttaactgtataataaaaattatttaagttctgtagcgttaggtcatgatttccagtg
 tagcaaaacgcgagggcacaattttttgtaaaaaggaacaattcttgcagtcagtaacactgagtaacataatgttgacc
 aaatgaagagcagagaatggggtctgcccacctggaggcagagccatgaatttggtactgaggcaaacaggagagctttgt
 gtgtgggtggaggcgcagggaggagcgcagggtctcatggaggcatgcgagggtctccagggtcaaacagacagaacct
 cagaactggtgtgaagcttagaagggcttgcaggaggtgtgctaaagcctcagtgatttgctatgggactggcaactaa
 caaagtaaacataaggggttagggtattttctttaaagtgaaaaaagtttccctaaactttcttaaagtgtaagctctgtt
 tcagcttaactcttgaaatgagaagtgtaactcgttgggtttcatgaactgcataatcacactacactgtttttctgtct
 acagatttgcttctctcagcacagcaattttatggacctggtacagttctctgtgattcttctcaggttaattttgtctat
 acagacttctctctgtttttttgatttaggcactagatccattaatagagtagtcagctcctcaaaagataaaaaaataatc
 ctgtgtttccacacagagtttaaaattcttcatatgtgttctgaaaagtttagatgtagcttacagtaataataaactcaa
 tctagcatgacaaagtagaaaaagtttgggtttgtgttctggcttccagcgggggttgcagttatctgtatataatt
 ttgcattgcattcctatgataaatttggctcttcatatggggagaatttggactaaaaaacacatttaataataatgggttac
 tgaattgagctacacatattgcttatgtgttcaagagatgattctgggtgattatccaaagttttaaagctgaaggtgat
 tttttagttgggggttttctttttatctccatagtttttttctgtatctaattttatcttttatgaaaaatactatagac
 aaataatttgaaaaagacaaaaaacgtgctgcaggtttttaaacaagaaagcagaataatgctgcttgcgcttccccacc
 ccaagctcccccaatttaaattgttttgggttttcttctctgtgtttacaagtttcttgcagtttaagttctctgtatct
 gtctactgtagaagtagggtatatactgtgttacaatggcctcccttagcatctccctcacaccacaaagcatacaaa
 agccccatctcttcattttaataccattgtatcataatgcatgggttaaatcagtgatccctatctatgttttagggcataa
 aactcttattcacagctgcatgaacaggtatcacactaccacatttttgggttagaaggatcattcttgccttatccta
 aatgtagtatttggcattgttggcttttaattgtttgattttctatgtgctacccctgatttatcactatactctt
 gctgagtacctttagaacagtaaaaaataaaaataaaagtagatttagctgtgacagaaactgagaatactccaca

FIG. 7

20/157

tgggtcacacacaggaagggtgctctctctgggtttgaatttctctcggagttgccacccttagaaccttcccacctcctgcc
 ccaagtgcaottgttctcaggcctgctgcacaaactgtcaattgggaaacttctcttgccctcctcctatattgaagtctc
 tggttcccagatcccattgtcttcaaccttccatccatccctcttttactggagcacattctctagatttatgagaaaa
 gatgctctgggagataaacttttctgtgagctctcctgaactgaaaaatattttattttatcctacaatttgattgaattctc
 gggggcagggtgtctatggctcacacctgtatttctagcactttggggagaccaaaagtggggtgggtcaacttgagctcaggagt
 tcgagaccagcctggacaacatggcaaaacctgtctctacaaaaaatcaaaaaattagctgggtgaggtggcgcaagcc
 tgtagtccagctaatcaggaggctgaggcaggagaatggcttgagcctaggaggcagaggttgcaagtggcgaggatg
 caccactgcaotccatcctgggtgacagagtgagaccctgtctcaaaaaaaagctctgactacagaaattc
 taggttcaaaatcattttccattagaatttagaagcgttatactatttttctctagtgtccagttataatttgagaagggc
 tagtgcctgttataatcccactctttgtatgggacccgttttctctcttaaaaaagaaagtataattgaagtataataa
 agttacagaaacatttgcaaatcctatgtgtgtgagctcagtgaaattatcacaaattgaacacttctagagatttgataa
 attatctcttgttccaccagaaacttcaggctcttct
 ctaatgatcccatccattgtctccattctcttgttcaactgagaacagagtttaattaccacactctctagaaagtctct
 ctctgtgtttccctcaotcacaacctcctgctctctcccaaggtaacctttacctgactctctcaaaactacattgca
 tttgtctatttttgaatgcatatgaattatt
 ccagagtcagtggtgtgtgtctctgggtcactgcaagcaaccgcccgtctccgggttcaagcaattctcctgctcagctc
 ccaagtagctgggattacaggcagctgccaccgcatggctaaatttttgtatttttagcagagacggggtttcaccag
 gtggccagggtgggtctggaactcctgacctcaagtgtccaccgctctgctcccaagtgctgggattacaagcat
 gaaccacccgctcagtgcaaaatgtatataaatgaaatcacaaataaaattttaaagtgtttgggttgatgaagaaaaa
 ctgaaaacttgtgggttgcaagttaaagttgtgtctcagaggttatagctgtaaatctgtggctctccaccttagctgcaac
 tagaattcacttgccgagcttttgatgtccaagccactctcagacactttgactcagaacctctagggttgcccccaaa
 catttgaaaaactctccaggtgattccattgtgcagctaaaggttgagaatcactgctttaaaagcacatttagaaaagaa
 aaaaactctctcaataactcagaactgcttaattaaacctaaagagtagaagaaagaaaaatataaataagagcaaaa
 attaatgtatctaaagggaagtagtacaactgagagaggtatctaaaaactgtcagttggatctttgaaagactaaagataat
 attgacaaacccctggcaagattaccaaaggaaaaagagagaaggtatattgggtatgtgagccaatgtcagaaataaaa
 aaggggggactctcccaaatcccgcaagcttttttctccagaaggggtatttgatgaacaaactttatgtgaattgaatt
 tgaaaatgaacacagatgggaatgggacattaaaagaaaaatagcgaacttgccaaaactgacacagaagtagtaaggagc
 ctgaatagttctgtatccatttttaaaacaacactaatttaagtttttgtcatagtgaaaactctaggccccagggtggctc
 catttggttaacataaacttttaggcggggtgcagtggtcgcagcctgtaatccagcactttgggagggccagggtgggtg
 gatcacgaggtcaggagatcgagacactcctgggtcaacaggtgaaacccactctcactaaaaatacaaaaaatttagcc
 aggcgggtggtgggcgctataatcccagctactcgggaggtcagacagagagaatggcgtgaaacctgggggagcagagc
 ttgctgtgagccagatcgccgcaactgcaactccaaactgggcaacagtcgagagactccatctcaaaaaacaaaaggaga
 atactgtgatcatcttaataatggaggagataaagcatttgataaaatttaataccactatttttaaaaaaattttct
 tagaaaaactagaaagagaactcttggaacactgataatggaacaaactaaaaagtcacacagcaaacactttcttaagt
 tgagaatgttagaagtttaggtttgggaacaagacagagatccactttccacagtttcaattactcagggtcaaccgtgg
 tccaaaaacattaaagggaataattccagaaataaaacatttataagttcttaattgtactccattctgagtagtgtgag
 aatgtcgtgcaactcctgctctgtccacactgggagtaggaatcatccttggtccagtgatccacgcccctatgtcacc
 caccacttagtgtcttggttatcagatcaaaaaactatagttatgtagggttcagttactatccatggtttccagcatcc
 actgggaggtgggaatgtaaccccacagattaaagcaggacactatattgcaaaatccaaaggataacttttggttctctc
 tctcttgactctctcggtgatctcaacaaactgactgcttcatgagcgagactctccagcgtctctgacccatgtctgc
 cctactgctgctctcttagcctctcatatgtcaactctttctatccagctctcactgttgattttttccacacacttga
 cctcggcgctctctcactctgtgttctcttgactggaactctgctctcccatctgagcagcagactgcccagaaagtgaac
 aattccactctctcactcgggtgttcccactcactaaactggcaacactaggtctcagctggttccaggtcactctcttgc
 tttcactcagcaggcactcccttaacaggcccccagagctgtggcgctcagtgcaactctccacctccactcccacccccc
 caacccctagatcacagctctcgctggacacttttcaactgactttggccctccagaattcccaaaaaagtcacagagaaga
 cctttaaagatcattgtatctagcccttcaacttccctgttgtatataaaataaaactcagcgtcctgggtgcaactggcgtc
 cgcttttccattatcagcttgcaactctctctctccacatttctgctaccaggacactctcctatgtctcaaaagacaaa
 gctcttccctcactcagcagatttgactatgcttccattttctggaaacactttctcactccctcactctgtatata
 ctcttcacacttctgttattttccagacaacaaacaaattctccaaattctctgacccaactcagtgctcaccagttcag
 ttccattcagacactaccatagttggagcaaatcctgcaaggttaggggtcagtcoccaagactgctccactcaga
 cactaaagcaaaatgggtccccaggctacactcagctgctactgggcccactacaaattcagggtctcgggtgactgctcac
 tccaagatttgcaaatgtgtagaacagctcacagaactcaggagagcactatcagtagaattacaggttttatatacag
 ggtacaactcaggaacccgcaaatggaagacatgaataggacaaggcatgggggtgggtgagcagcactctctcagdcg

FIG. 7

21/157

tctctggacaagcctcctcctgcacgtggatattcactatcctggagctcccccaagcctcatcattctgagttt
 ttattgaggtttcattacatagcgatgagtgaaatcactttctattgggtgactgaactcatctccagcccttcc
 cctctctggaggttgggggtaggagctgaaagtctaacctctaaccacctagttgggtttctggtgacccagccccc
 catctgaggtcctcttaggccccctctcccatgagtcattcattagcacacaaaagacactcctgtccctctgaaaattcc
 aaggggttaagctaaattgtctccctctgttccctcactcctgtagaactgtatttttcatcgtggtattcatcacaat
 ttgaagaatatattcataatgttttaactgtcttaatatctgggtcggtagcccaagcttcaggggagcagtgcaactgtctct
 cgttcactcctatacacacacagcttcagactctagtacataatctagtaagatttttggggaagttaaacttagactct
 tcagggttaagtgcagagtgaggaggaaacctagggcaaccaactctcagggaaaaaagcttgtgtcaatgggtgtctttaa
 agggatgtaggggtctgattgttagctccagtagacatccttaattcttggagcatgtgtatggaatagatggccctgc
 tgccaagtacgtacagcctggtgaggaccggccactctcactggacatcagtcctcctgggagccgaatgacactatcag
 aataagtgcaggaagtacttttcaggcccatggccctctgtgtagtaaacaggacagatggcattttcatcctctcc
 cttgtgggggttgaactctcgaaaaagaatcggaaactttgggtttctagacgatgtttactaaactagcgatgataaac
 ttttagctaaagttagagttactttcatggcatggggaagaaaattttccaggacaagtttaacctgtctaaaagaa
 ttattttagcttt
 tcatatagccactgcagcctcaaacctcctgagctgggtcgaagcagtcctcctcactcagtcctcccaagtagagctggga
 ctgagggcatgcaacctcatgcccagctaatatt
 ttaaacctctggcctcaagcagtcctcccatcctcagcctcccaaggtctggggtatccagcgtgagccactgctccag
 cctgggtccaaaagaatttaaaatggagcctataaggaagggggaaggtatccagataatcctaagcactctgaagaat
 ggctcatataatatatggaaaaaagaaacaaaagtcggggggaatctcgatgattcaaatgagaaaaatcaatgtaaagtga
 gttagaaaaaatatttccagactcggggcacagaaaagtaaaactctaggtcctgacacagctctgttgagactgtgtccta
 atgagatgcttttccagcctcaggaaaggacccctccttaaagaggggaataagagactctcctctgtctgtctgtag
 cttctgttagctggagcagcaagtagaaggagctggggaactgggtgtgtgcacagatcttcaactgaatacagaagggcc
 ctctcactctccctgaggtcatagaggggaagcaggcacagactgtattacgggtggacccagagaggtttctgagtga
 cagggtctcctcctctcagcagcagggcaggtcttgaccaacttagaagcttgccctggcctcatgctctcctctct
 tgtggatttgcctcctcttatgttttatcttagaataaaaaataacaaaaatcaggattatgctgtggaagccaaagaa
 actgagaaggtttattctcagactctcttactgtgtctcctgtctcctcactctcattgttaggtggcccccaggac
 atttccatctatctgtattcactctcttggtagtgaggccaggtttattgcttgaaataccatatatattccttatgct
 tccaaatgccatgccagccagacactggcagccagcctccaggttagaataatcggccaggacatctccctgggctgt
 ccggttagatatctcagggttctgcatgtctaaaaatggatctactgatatgtccagattccaaagtgttttcaacctct
 gcttccacagctctgggttgagccaccatcatctcacttagactgtgtgtcctcctgataatgccattgcaacctctaagt
 gttttctcagcaaacggttaaatcacacacactgtgtttaaataacctgcactggcttttcccaatggccctcataatgtgtg
 cagtgccccaatgatctgaccttccacctcaccctctctctgctcttgtctctggccttacacaccccactccaaaca
 tgctggcctcctgaggtcctcccaagcctccaggcacggccttgctcgggctttgcaactgactgctcttttccaga
 cgcaacttttgagggtcctcatgagccacctcccatcactctccaagtctttgttcaaatgtcacctcttgagaagaa
 ggccacacactctgacctcctgtgtggagtgccactgggtactcctgaaccccttgctcttagcatctttttattct
 tagcaacttacctaccttttgcatgtctgttaatttgcctatttacttactgtttgtctccactgtcggaatgtaaat
 tccatgggacagtaattcttctgtttatcagtggtttccatgagcctagaacaggacctgacccaattaaagtattg
 attaaaaaagacaatgtagccagtcacgtgagggcctgtaattcttggcaaggcaggcagatcaactgaggttaggagtt
 gagaccagactggccaaagctgggtgaaacctcgtctccatataaaatacaaaaattagctgggtgtgtatggccgcacacctg
 tagtccaagctactcaggagctaaaaagagggaatagaatcgctgaacctgggaggcagagttgcagtgagctgaga
 tctgctcactgcactcagcctcgggtgacagcgagactgtgtctcaaaaaaagacaagttagactgtcgaagaagt
 ggcagctgtctgatagcagagaggccacagatgaaggctttcccaaaagtgcactcctgggggggttggaaggcaact
 ctattgaattgctgtgtgtagggcaccacttggaggagctgcacaccttgtaaccttagggaggagcagtgagcct
 gagataagggcagttcagtcactcagtcactcttagagtcctctctccagccagccaggagggtgtgtgtgtgtcactg
 agttgggtgcactgaggtgcatctcatgaagctgaccccatctacaaaacgcagagctcagccagcaagagagtcactgt
 atgacactggcagagagccttatgtctgtatgactctggaaaaataacctgtagctagcccttttacaagttaatt
 gagttaaaatgcagtttatagcagttatctatttttacttgttaagcttatctctcttttaacggcaagtgcagctgtt
 ctctttggcatcactgtagcttttaaatcttttccacacctgggtttatacatgagagagcattgaaatccttgcaacta
 ttaaggccactgggagttattttctcttaatttttaagaaggttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
 atttacttaaaaaatattgtatttgaataaattgaataatacactaataatacacaactttctctgtgtgtgtgtgtgt
 ttatgaaaagtgtattgttacttttaataaacttaacttttccactctcttagattttaccagtaaatatctgctaacttgaa
 taacagcagtcactaatatttgttatataatcttagaattattttcttatttgggataagccactgtgggcaactctg
 atgaaggccaaagtgaaggtcagggctctctctaataccacatgggaagcagtcagctcagggcctgggttaggaagcc

FIG. 7

22/157

tggcttacccaagttgtctcactcattacctcatgagccggcacattgtgtgacaaagcatcagaagggagcccgagg
tggcatgtgggagctgttttctactgcctgggctgaggtgtgggcaacgggtgggagctgattctgtggggctgggataata
actcagtaactttctcttctcacagttgtttctctctcttgctcctgtgtgctgtgtgtgttggaaagatcaacaaagtt
gttgggctccagacgttagagaggttaagcttcagtggttaagattaaagaatccctgggaagagcttttttctctctt
tctcttaagcaagttgggttttagctatttagtgataatggacagacagacatctccacgggaaggaataatgcagccctt
tctgcagtcctaggttctgcagcatatggtcagtagacttagatgagtgctacataaagtgaattagttctgcaccaca
gtgtcttaacctccgaaaaggaaactctgttattgacttggaaatgccagttttatccttcagaccagaacccaagtcc
gactgagggccagagatccaagggcgtatttagcacagccacgcgtgattaggcacccgcttcctcatgagaacttgtcatgcy
agagggaaatcagccatttaagcatctcaaaaattttctattatccaaggaagataaattgtgtgtgtagtttaagtttt
aacttgagttttatttttaataaagttcttgatttgtttccagagcttcattccaagaagtaataagcaaatatttacagct
gacaaatttaataataaacttgggtgcaaatataacataataatgcccatcaatcaatgagtggtgaataagaaatg
tggggtgtgtatgtgcgtgtgtatgtatatatatatttatatacatattgcagcaacctggatggaaattggagaccgtt
attctaagtgaataaattcaggaatggaaaaccacacttcatatatctcacttacaagtgggaactaagccctgagga
gcaaggcataaagaatgatacagctggattttgagcactcgggggaaagagtggggaggggaggggataaaaactata
cctcgggtgcaggtgtacactgctcgggtgataggtgcaaccagaatctcagaaatcaccatgaagaactttatcatgtaa
ccaaacacccactgttcccccacaaacctattgaaataataataataataataataataataataataataataata
aataaccacccattaaataaattacaacataaaaaacttgtgtattgcaacttctcctgcacatctactttgoccttcat
ttagtcctgcctgtggcatttgttggcatttgtgtctaattgaaatgctcagctgtatataagatctagacaaagaagctgt
ctggggttttttggatctactcagctcttttctataaaaaatgtatttatagataagatttagaacagaagggaggt
atgtaaaatcacaaactcacaaagcactcagcactgtttagataggtattcatgtctgaataggggttaagacaggtccagga
tatgggctctgtgttggggcagctagccacttttctgacccacataaagataggttatcatcagaggggaattgtgca
cactctttatttattttttaataagtggtgaatacatattcaaaaggttttttagataaactgtatctagctatacttt
aaaagcctaagtaactcagtgatgttctgtaaaaaattatattttcaatttaatttaattttatctcacaatgga
gtaaagaaagagtaagaagccacagtaaacaggttatcaagaaggagtgaaanaacttttgagtgctactctctccaaga
aggacatagaccagggaggttccaaggttgagatgagcttatggagatttagatagcctgttctcctcagtgcttat
gagaataaattgtttctaaaactctgtgaagcaaaaggacagatggaaagcttgaaagagtcgaagttagatttgattatca
taactcttggtaccacaaactgacaaaggttagccacaaatacaaaagggcagctgcattagtgaaatgtatgtgcagcagtc
taattgcagtgtaatagaagagtcacttaataccagaaaaactattagaatatcatgtctcaaaaaaagaatatatcatg
tcaatggggaaaaaataacacctgaattctctcctagtgggccacagaatactttcgtcatttatgaaactcaatatccat
tctcaatttgaaaacaaaactcacagtaaaatacaaaaactccttttaacgctatgaagaatatctggataaacacagcatc
actcatgataaaaacaaaacacccagaggttgtatccaggctagttagttttaacaagagctgacatctgagcatttaa
aaattccaggcactgagtttctaagtggttttacgtagattaaactaatccttatctcaaacctatgaataaggtattgtcc
ctcgttatgttttatggatgaggggactcagggcactcagaaaggtgaagtgacttacccacatagtttaactaaagggcaga
cgaggctttaactagggacttctgtcttcacagcccaagttcgttatctcgtgtaatctcagctcattcagctttgtg
ctagatgttctagactgtgcactaagtgaaggggcactaagaagtggttgttactattggaagggaacaaaattattatt
ttcataataagttattaaaacaatgtgtagagatttgctgcataaaaaagaaggcttaagtatatcaaaaaataaattg
aataacacagaagctaaacaatacaactggaaatttgtttgcagtggtgcatgagagaagtaagatgcacagccctacaggaa
aaagattaaaaatatatatgtgttttagtatggagttggtcatctaataaaactagcttttaccctgtgtatttgaatt
tctctgttaactgtgaagctgaccattcagaataattttcatcatttttaaaatgtcagccattatttattccctctcaga
attgtagaagcaatgatttgatgtttttttaaacaataagtagtagcggctgagctcattttagtcctgtgtacacacc
agccccggttttaagtgcttttacacatgttgctcctgaacctgtcttatgttgacccagcacttaagcacagacc
caaggagcatgtctgcataatattagtcagagtaaatggacctacagctttaaaaaattgaagaattgttggtatttaagaa
aatattataaaaaatgcctctcagacagggccatctatccaggccaaagtattgtctctcagacttaactgaaggtccaagga
ggcagcttggaagtgaagccactctgtatcccaaggcataaactcaaaagcttcagcatcatgccagaaagatctctaata
tcttaatgtatgtgaacggtgttccacatgtgctagggcgtgtgctgtatctcattttatctcctgtgacaaact
tatgggctagggaacacatttcagagggaggacacggatgccagagacttcaaacaccttcttaaggtcacacagctgt
agactttggagcagggagttcaacatcaagcctccctaaatcccgagccctgttgtctcctcactctgcacaaattccac
atttgtttgttttagctttacacactctcgaaggagggttctttagcttctcctgtcctgtataatggaggtcagg
ttttgtgtcctgtgaaactccttggtgcagttttagtaggaagaaagcaagcctgacttttagtagtagaaatacagga
gataggggccacacattttatgtatttgacaagtgcagctcatagcattaatccacagctatttatttttattcttat
ttcttattcttactttaagttctgggatacatgtacagaaacattagattttgttacctaggtatacatgtgcggtgtg
ttggccgacactatcagcctgttatctaggtttaagccctgcagcttgggtattgtcctaatgctcctccct
tgcctccacccctgacagggccaggtgtgtatttcccttccctgtgcctgtgttctcatggttcaactcccactt

FIG. 7

23/157

aagagtgaggacatggagatatttgggtttctgttctgtgttggtttgctgagactgatggcttccagcttcatccgtgt
ccctgc aaaggatagaaactatttattttaggcgcagatagattccatgggtgatatgtgccaatgtttctttatcca
gtctatcattgatgggcatttgggttggttctaagctcttgcctactgtaaatagtgctgcaataaacatcgtacgtta
atgtatctctgacaggggtctaaaaagtagtatttggacttgatattgtttgaatcaggaggagatgaggtttttaa
cattatctaatatcaaggtcttaaaaaataaaggacaatgggttatcacatccacaaaaatttagtaatccctaatg
agcattaaactaaaaacccctcagctactcattccatataatggctttgccaatattattttagcaactctctcgagag
atgcaacagatggccagcgctccctttgacctgtataatgtcgccctggaaacagatgaggagcctcctgatctattgg
ggggagtataaagggtgagaatgtgactcagaagtcctataaacttgactttttaaactaggctcctaagctcgggaaa
ccagagagagacaaaagccctattccattacctccttttttctctctttctttttaacaataagctaaacccgaagt
ctgtgggtactctttgtctttgttttcaactctttgtttcatgtcgtgagcagctctgtgtcaggggattttaaactc
taaatgattttgtctgtctttcttttaagtattctgtggcatacaaaactttatttggaaagagctacattgaacatcga
atcattgtttttttgtttgtttttttagacagagcttactgtgtcacacagctggagcgcagtggtgatctc
cagctcactgcgaatctacctccgggttcaagcgatctcttgcgccagcctcctgagttgctgggattaccgatgtg
caaccacgctccagctgattttgtatttttagtgagatgggtttccacatgttggccaggctgtctcgaaactcctga
cttcaagcaatcccccactcggcctcccaagtgctgggattacaggtgtgaaccacgcaacccgctattttctct
tatattttggggaatttttaagcagcaaaaatgataatacactgtgtatataactctgtggaactttatcacccccc
ctccccaaaatgatgaagaaaatggtaagactttctccttgggcaaaaatgatggaatattaaatagctgggagaat
gctatgtattctgaataaaaactacttacttaaatgttactgagctcacaggaagtaggtaaaactccagggaaact
ctccttccctacccctaaaagaaacaaacataaaacccaaacaggaacatgattgtcctaagtggttccagatgggt
caacaacgccccactgaatagatggaaaatcctctctatagaaaaacatcttcaatcagcgctgttggaagtccacagat
aagatgaagtagtaataagatcacaggggacaagggcaagcgccataaaatgaagtagcagaaaacacagatgattgat
gtggacaggactccaaaatgtagaattgaaatctaaagaataataataaactgtgtatgaaatgtttaagaataaaa
atgaaatcataaagatgagagtcaagtattggccaagcagatctgaagaagaacaaaatggaaactctagaaatgaaa
atactcttggaaattttttaataataataaacttatactgttaacacgcatatcagacataagagaatttacaagt
ggaagtgaatctgaaaaaagtcagagaatcgagctcagagagacaaggagatgaaaagtacaatagatactaagaat
tgaggatagaatgagaagatccaactcatatatttccagaatccagagaacaaaggagggcaatattcagaataatg
agcttccagaactgatggaaaacatgaatatcacagattccagaagcaaaaacatctgaagcatgattaaagaaa
gggaaggcaggaagggaagaaaacaagtgggatttctcataatgaaatgtataactccaaagacaaagaaacatctta
aaatcagccagagaagaaatgatacatatacaactcggagaaactgaacgttagtgactgttgatggaagtataaagt
ggagccactgtggaaaacagtaccactattccccccagaatataaagtgaatattatattgatctggcaactctgggt
tatacatagaagaactgaagcagagtctcaaaagacatattgaacatcagtggtcttggccagcatattcacagtgc
aaaatggaagcaggcccaagtgctccatggatggataaattggataagcacaacgtgggtctatgcatccaatggaat
ccagccttccaaaggaagcaaatctgacccatgctacaatgtgatgaaccttgaagacatactaagtgaaatctccag
taacaaaaaaaataaacatttgtatgtatcaccttttaagagatacacttgtgttagtcaagctcataaagacagga
aacagtggttccacaggggtctgggagaaggggaaaaatggggagttagtggttaatgggtacgaagttagaacact
ttaaatggctcaagatggttaaatgttatgtgtattttatcacaaattttgttaaaaatgggaaaaagagggcgggt
gctaacaacctgtaatccgagcactttggggagggcggagggcgggatcacttgaggtcaggagtccaagaccagcctggc
caacatggtgaaaccccatctctacataaaaaatcagaaaatttagccagggcatgtggcacacatctgttaatccag
ctgggaggtcagggcatgagaatcatttgaacctagaaggcagagtccacagtggagattgcaccactgtactccag
cctggggcagagacaagagctctgtcgcaaaaaaaataaaaaaaaggggaaaagtacacctgc aaagggaatg
acaagctgcacgcgttaactctcagaagtcagagaagagtgagagaatgtcttgaagtagtgaagaacacctccccac
agaatagtggtgccagtagaacctctcgagacaagagatgaaataaagatatttcaaatagaatgaaatcagagaat
ttgccaccagcagacctgcaaaaaagaaatgtctgaaggatgtgtcgttaaaatgaaggaataatccccccagga
gagatgagacgttaagaaagaatagtgagtaataaaaatgatacatatagaatctaaacaaacatttctcgtgtaaa
atatttgggttataaaaaatagcaagcctaaagtattgaacaaatgatataatgtcagcagcatatgacacagtgat
gtcctctaaaaacattttcttgattttcttcaggacacccacacattgcctctctagctgttcttgctcagtggtcttgc
tcagctctcttcatcttttgacagtggtgtgctgcagggcctgctttgtccctctttgtgctcttttgacacactct
tcagtgatcttcatctatcacaggttttgagatccactgatcctctcactggactcgtccatgagcctccaacacaa
cagttgctcactactcagtgctcccaacttgatatttcaacatggccatggccaaactgagcttttggtattttctcaga
cttgctcctctgacataccttccgctcttcagacataaacctccatcacctagttgctcaggccaaaatcgtggaat
tgttcttgactcattctgtgatgctgtctcattggacagcagctctatcagcaaatcttgactcgttattc aaagt
tttccagccactgtgtcccatctctgcaactattcccacttgatctaggccacgcgtgtctcactcgtgattcttgca

FIG. 7

24/157

gttacttcataactgatctccctgacctcgcattccaccatctagtcagttccctgtgcccacagcagagtggtactgga
aattattttcttcagccaaactgaagaaactgatgacattttgtgggtatagtcagcagaattctaaaggtggccctcaag
atttccatcccacgtttattcagtcacacactactctaggtactgctagagggattttgcagatgtaatgaaacagggga
gttgactttaagatacagaaaagttctctgggtgggtggcagaaaagaggaagccagagagatgggaaacatgagaaggact
tgacacacccatccctgggtttagagatggagtgcccacagaaagttttttgtttttgtgagacagagctctggctgtgtcac
ccacactggagtcagtggaatgatcattggctcactgtagcctcaaacctctctgggtcaagggatccctccctgggtgc
tcccaagtgctgggattcacaggtgtgagccactgcgttaagccctgaatgcatttcaaagcagagctctcccagagctct
ccaaactaagagtcoccatcagctgacaacttgacttcagcctgagagacctcaggagatagctgataatccccagaact
catggagaacatgaacagacctgtctctctcggcagagaaacctacagagcttgctgggactctgaccttaaaagctgt
gagataaatagtggtgtgtgttttaagcaactaagtttgagtagttgtttacactaatatagctgggtggtgaggtc
acacctttaatcccagcactttgggaggtcaggtgggaggtatcgcttgagcccagaggtcaagacctgctgggacac
atagtgagaccccatttctattctttaaaaaaaaaaaggaaggaagaaaaagagaagaaactaatataggggacaaat
ctcatctaaaggtgcaacagaagatgaaaacataatttcagagccaaatgcactcagagagtagacaaagaatctact
tggtcagccaaacccagcttttacaagcttttaccatccagcaggcttgctactagttgtctaaaggacagacaaaataga
atggtttttaagttcccatgatctcgtatctgtgtttgagccaggtatgcagaaactatctgacaaactctaccaagcact
accactggcatctcgtgtgacctacgttatgaatttgaatttcaaatcgagccaaagtcagcatccagaagaaaccca
cttgccgtggccctctccctacagttcaagcagccctctggaagggagctcgagagacctcagtggaaggtgagcaagcag
accacccgtgtcactgggtgctcagagagtagtttagctctctccctgaccttttcatcttcttaaaaaactgcaatgtac
cttttttatataactgaaatttgatgggtcttataaagcacaaatgacogttaagttgtccgtatacattaaagcttgagta
tcacttaaggttaattttgtttctctctgcagactgttcccaaaacccatgtcactggagccggtgttttggaacaaaagcc
gctgtccctctctgtgtttgtgaggtccctcgaggtcgtccctccctcctgggagtcaggtgagtagatgcg
gtccagcgaagacacactctaaagcatgtgagggagctaaagcatgggatacttcccttctctagagaaacaggtttataaa
gggtgatacccaacagtgctcgtcccatgcatggcagaggttcaagtagataaacccatttcaactatgcccacaaactg
ttttcccgaggtcatcgtccctctctgggggagctcccgagacctctcacccgtctctggcctcccagacctctcagtcg
tcgtggtatgctctgaccacacacttctttgtgtgtcagggtttctgggtcccccttgctctcgtgtcttctccctgtct
cttttgccctgtctcgtccctgaatctgggtgtgctccctccacccgtcagttcttccctcagttgctctccactcca
aacctccagccatctccctgggtgggtgtttcttcaattgggtatttctagccctccctgcaatttctttgtcttttcaac
tgttttaaatataaggcacaacacacacagataacataaaggaaacgtctgaaatccacgcacagtcctgctatcccaca
accacacaaattttcatgttctgtgttcttctccagccctggccagcaatttagttgaaatgatgtatagatctaatttt
tattctgctttttcataataactatttaaacattttatagcctactacataatattttaaccaatgcatgtgacactc
tccatcaaatatagtaataatgattatataaaactctattttgaagacagaaaatatggaatacagttttgaaggtatct
cgttaattacagagtgctttggaggatgctcgaatgtgatcatttagacaacacttccagttgagatgattctggagtc
gggttaccttagccagagagcaccctgtgtctgtgtggggaaggctgcccaggatactggtccagggccctgagctccactg
tccctcagtgcaacctcagacccagagaagatctccagccacagctttagaatgggttaggatctatagtatcttccagagg
gaggtcttgcacatcggtcggccaggattctatgtcttctcactcagctagggttctgtgagttgagcccactaactcg
gcttgccctgtgttgaggacaggacaactctctcgtctgctcactcaagtgaggtttgtgtctcctgccaccgcccagcc
agcaaacagtggtgtgtgctcgtggggcaggggtgggggtggggcggaagggtgacaggtgtgcagccctgtagcctt
tctctgctccagcctttgctgacatgccccacccacactgaaacacactaataaaggctctgctcaatgcttgagacccccc
aacagttggggcgtctctgtgctccctctctgtacccctgttttaagtggtgcccctttctaggctcccagagcattccc
cactgtgttccaaagcacttaacagctgtctacccccccccactcccttcacagagctcaaacaggaataacacgtctc
cactaacacaaactcctggcaggtctcagcagagctccagacctcgcgtggccatttcccacgtatattaggaataccagct
ccccagaaacgtgcttccctaatggatgtcagattactcgttatcaatacagggaagaaatgggattgcaagctcc
caaagaaagtgggggcgggcagctgtgctgtgtgagataaagggttagttgctatgtcgtgggtgtgctggcgccagatact
aagactttatgcatgcccaggtactaagactttgtcagttcactcattgattcttccactagccctgtactactattatc
ccacttttcaaatgaggaagctgaggtgagggaggttaagaaacactgtctatttccaaagcctcctggcccagagctgc
cctgctgtatcgtgtgtatggaatggaagtgcactaatgctgactgagtgataggggtggaatctggaatctcgtcataaag
gtcgtgctattttctcggtcacttttgagggaaggcaggagaattggggacttagcacagctctacaactctcaactctc
gttcccccacccaaagcagagcactgtcgtgtgtgtgtgctgaaagcaaacctgcaactttgtgttaaaagtagggaa
aggagaagaatgtgtattttgtatttgtaataagtaaccccaagaaatacaataaaatagtgaaagcgtttaccagat
ggaggtggaagcggaggggggtcaggttgggaaggaagacttttccctgtcattgtgtaacttttgggggtttttgtgt
ctgggtttttgtttgtttttgtccactagggattgaaaaataaattgttttaatttttttaaaacaaaataaaagc
cagctgctccagcagatagttctctcgtgtgtgtctcccaaggtctctgctgtgcccagcctcgtgtggaactttctcagc
agatgccgatagtgtaagaaggagaagtcaggagccgtgagaacccgaagcagcagccctcgcagcctgggaactgc

FIG. 7

25/157

ttctgatgctggggccagggaactctccacgcagcagtagtgagtgggcacacagagacatctgcagggaagggcgtg
ccgggaataaggctctggcggtgcgggacagcagactggaacacctcaagacatctgactcaacctgagtcacaaagctt
cttttaagcgtttctcccatcgtgttccagatcttaaccttttaactttgcataggaaatacttgatttaattacaggtc
cagggaatagactgatggttgcctggaggagggcgtgtaggcaggtgagagaactagggaatgacatccaggttcaactgtg
gaaaactcgtttcttgggactgtctcaactcgtgcaaaaaaacaagaatggaggtgttttacaatgagacatctgctcatcagttg
ttcttgaaatggctcttttaaaaactcgtcagatgaattaaacttgtttcaatctgaaagcctgtatcttttttaaaagt
gtgctatttatctctgacagatttaggcgaattatctctccaggagtagactttttctagttagaataataatg
gtccatctcttttgatcatataagctaggtatagaagggggcttttaaatgtcaaggtcagcagtgcttacttgaat
gtaaactgtgatataagtgagtttcttatagtaacttgattaaattagcttaatactttgaaactctctctcccttt
tctctgtgctgctctctctctcatcctccctccctctcacacatcacacacaaacacatcacacacacaa
gtgcttagactttaaatagattagccaattgggaaagttagtaagcctaagtttttacaataatgcattctcatactctt
taaaatttaaatagctaccattggcaactctgctttttttctaaaaactgatttgcagccaggaaagaattttctccac
aggaaacttgatctagcagcaggagtagagggaaagcagaataagtaagtaactgtgaaagctcctgctttttatcaaa
aagcgaactgtcaagaaagggcccccctgcccccacccctgttcacccctagggctgataagcgtcagagggaaaggaact
attcatgtcagcgtctctctgagcagaaaagagcactcagagcacttgggacccctggtacagagagatcctgtgctgct
gcagcctcctctgaacttgtggttcattctcaggtcgggtggactcagatgcaggaaggaaggagcagccctacttgtc
ggcagaagctgcccaagctgggaaagctgttttgctctctcccccagggaggggctcgaccaccacccctcctct
ctcagccaaaggttggtggtctgaggaggggcagcaattgctgacaagtagtaaaagcacatggaaaaaattggacagga
gggaaactctgccaaattggaaaaatggcacaatttaagagggtgggacagctccctgctcctctcagcagggcactgct
tgaaatttgtgttttccccatttatgtctctgttatctgtgcattatgcagcagctccacagagctctctctgctct
aaaaactgggattctctggcattaccctattgggattggacgctggacagcaatgctcgagatttgtgaatttggagagata
ctcaaaagagctaaaaactgcagcattttacctttaatgtgcgtgctacagagagagatttgtctctccccacactaa
ccccctccctagaagaattgctggaaaagtgttttcaaggaatttgaacataaaaacattctgatgcacagaacac
ctctactttgagatcaactctctataagcgtctctttcacattatgcttaagagacagagctctagaagaagccctc
ctctcatgagctccccactccctgtcacagacagccacacccatggcgctcgagtggaactggcccttaattccccac
gccccccagcagggccaaaaggaggccctcggttatgtcctctctcaaggaagatcctttgtttgttcaaaggac
agttttctcaggccgaaagaagtctctccccattgttagtctctatgcaattcatgcaactgacccacgcatc
tgtgtatgtctatttttttctcaaaagtattgttttttaaaagggaaggagagcaagtggaatttcatctgtc
tcagcgctgggggaagcgcgtgataccccctgctctctcttgcacgcgacagcaaaagcatttctccggctcagggc
gaaaaagggaattggcaggaattgaaggcgctggctcggagctggtttccaaggaagatttgttgtctctggcggt
ttgcgctgcacaagagagcctgtataataaattaaaatagtcagaacacacactgacttgcacttgatacataactacag
tagtgtccagaattgtcagacattcgggagtgtacataaaacagaaaactctctatgttttttataaataaaatg
tctgagttttcaactaagattgtttttgtgcatattcgtgatatcaggttctgcgcaggcccgatcaatgaaataacaaa
cccaagaaacgcataccccatttgttgatgtgtcagatgactctgggcacacattaggaatttgtttctaaaaaggagctact
tctcagagtgcatcagataaaattcaggcagggaattcgaaacgcagccgctggaggagactcagggaagcagagggctccgt
ccgtgcccccttggccctgcaagcacatcatgacctttctggcagcctctggtgctcgggtagtgaaggatgacaa
cttgtctcgagaattgtttctctcttagcttttaagttcaagaactaacctgtagcaatcagacttccaaaagggggtct
ccattttttgtagttttctctaaatttttaagacatttcttgaaatcagtttatatactgaaactgggggtgggag
taggggctagttttgttgataaaatagttccatttccccgtggagaaatttgacataccctgagatcctgtgtgctcctg
cactcctgcacacacagctggggaagcctgtgctccccctgtggagagaaggcaccccgagatccccgagactaac
cggaggaaggcagctcctggacagaagactgtcagcagaaggaaagtactggaactaccctgggtagtgctgcccattca
agactgcagacactgggaaataaaaagcaggggacgtcgtggtgggaaggagcattttaacttccagtgcaaatcctg
ctcttttgatttaagtggtgtactggggccaggggctgattcaactctgggagtagtggtgggttttcatgaacatct
ttgatcttccattctatttattcatccatcaaacagatttgctaacacataacttaagctaatgtctagggtag
tgactgagatgtaaaaatagattttaagattaaaaaaaaaaccaagctctcacaccctgtcatccccagagtagcttctc
ttgtgtgtgtttctgtgagaatttggccatcctgaggacacacagccaggcaggcagggctcctggcctcagggcattgccc
tgctacactctgaaattgttaccacttgacaaacacttggtccagcactcgggtgtttctctagatagcaggccacac
caagatatttgcccttgatgagtagtaaaacactgctcacaaggagatgtttgaaattggagaggaataattggcaact
tcactcttttaaaggcagatataggaaattgatttttaagtaactgaatttgtgcacaaaactctcaaacactagtgaagct
gtttcgttgaaactaatattcttgctctggaaatgtttttgtttatagttttatgagttctgttttttgatttgaatcaa
gcttagtttttaattatgtataatttagcatctattacactcatgtaaaattggagtagatttgtaaaactatttcattg
cgggattgtgggtgttatacatcatttaggcgtcaatttttggattttttgttattgatttgaataaacagacttaatt
aagcaggaacaagagacataaggagagctgtgcatctttaaacacaaattgtaagaacttgataataacaaaagtaaat

FIG. 7

26/157

actataatacaaaacttcctctgaaataaaagtagatctggtaaaaatgtggccttttgtctctgagtggttccatttttgat
ttgcatgttttggcctatatattacattttgggtcattagaaaccttaagggaataaaagcggtttaccagttttcagact
cgctcagaggaggacacttgacagttaacggaggttgagggtgaaaacccctggcagggtacctcagacagccctccaggcag
cttgtgagccacagaccttccacacagcctttgtgtgctacaggggcagctgggcccgtgggggcagggcagagatctgag
cagtgagggtatggacagccgtctcactgtctccactaaccacaccactgtctccagtggtgctgttatccctccactttgc
gctctcctccagggtcctgaattgtctctagggtgtctccaaactgtctaaatctcagggtgtgtgacagggtcctcagccg
tggccagtggttttctcggccttgaagttttagggaagttactccatttggcaaggccttggttgggggcccctcctgtgtg
tccctgcccccttttattttggctcctcgtccctcagaatgggtcagagtcagtgtattcagtgccagagtgccagattcccggtg
ctagcagttccttccagaaggatgaggtcattgaaagaaggaggaaacacaagacgggtcactcctgtagggaagatggc
agggctactagggccaggccctgaggtgaaggcaggagcctggaggaggcctgctaccagagctcaggtctggcttaggg
agggaaatggctggactagtggctgcaaaacaggttgagaaggaccttgaacctcttccaggagggttatgagagccacc
aacagtaacaggatagaaaacccctgtgattttgggtgtactgtggcagagacaatgctgtgtttaccaaacctactggggcc
acagctcatcatttccccagcctcctcgagttagggtgggaccagctaacgcatggcccaaggaaaggttggaaagaagtg
atgggcccacatttccaaacgtggccacacaaatgctcccacaggggactctcctacggatcacagttaagaactggagg
actgtgggtccctgagtcaccccttagagggccggcctgagtcagtaagagacttctgtggcctttgcatcactggaaag
caactcactctcgtgatccactgagatttggggctttgttacaactagttagctttcttatttccactaatatacaaat
cagcactttgaagtgccagtgccacattgaaccaaacccttaagtaacatggcactggctcagcgtgtggccaggtggcgagc
agtgatagaacaggtgacgtggcaggacaactgaggtccttggcaggcggagcaaaacatctggtaaggccgtttcact
cgataactcggaaaacaaactccatgtctgtgacttttggacttgggggaataaaacaacaacaacaacaacaatatg
tgttgattgttactggctgcttttggcgaggctacaggaataaaagttactcaagaagaattagccagtttacaagc
agaagggaataaagatgaaagaattgttaaccaggcccttgcatttctggaaaagccaaactgctctagcttccaaaccagt
agaggtgacaggtgagcaacaaggcccccactgagacttagtgaacacagtgactgggcccctatgggtgacatccaatgaa
gggtgtggctcttccactggagctggaagactgcaaggccacagcttggaggaggggcagctgggttaagggaagcaaga
aaaagctgacatgagaactgtttccagggaacaaacatgaatgtagtcactgacccctggacccctggcctggaaataataa
gatgtgttaaatttaagaataacatagcaaaagccttaaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaa
gtcttaataccacttccaggcaggaaatgggtgtgaaagctgtgtccagcgagggtcttacaagctcagatgtgggccc
aagagcgaactgggctgagcaggactccccagccaggacagggttgcagaaacacccagcagagca
gcaattgctgcatctgcccgcctccttccctgggtaggagggtcctaggagggtgtcctgctgttctcactcactcctgc
tgggtgggaggtggattaagctgtcttgttaattcaggattgtggacttcaaaagaccacctggaggaggtcgcactcacc
tggaggtctgggggggtgtgacctgctattgtccccctaggagagggtgctgtgacccatgtaggagagggaagcaca
aactgggtttgggtgatcagaatagcaaacctgggcccagatgcttaccacacatttccctcctgagtcacagctctcgac
tgcattggcctgggcccgtggcatctacccagatgcttgggaccagcctcaccatgatgtgtcatttcaaggccattc
acagtgacagtgctcctccactgtctccactcaaccttgaaggcttatgggtgaagatgggttaatacaacaggattgaagga
gactgggcccctgaaatgactttgtggaaacagacactataggccaaaggtagtagaattagggtatctgtttttctgtctg
ttaagtgaattaatattctctagatttatccactagttattcaacaaatttttatttgagcacatcagggtcagggca
ttgtgtcaggatcagagattttgcaatgaacaaacagaaaactccatgctcctccaggggagttgataataaaatgtcag
ataagctctgtgaagaagggttaaagtaggggttaggggtagaaagtgtctggggggatgtgctacttaacatagtggtggcca
gggggaagccattctgctaaggtgacagctgacagcaaacgatgagtgaggggggcagggaaggaaatgtggaggtctgggaga
agagcattccaaacagggtatggcagggtgagaggccctgaggggggcagcagctgctgttaggctcctgtgatatctgtgag
aatgaaaaagtcatttccagctgtgttgccgtgggaacctaaccaggttataaaacacaaagtaacttgagagctgggtttaa
tatacagccagctttccagcagtgcaactactgtgtacatcaagggaataactgaaacttcttcttaaaacttatcat
cagctgggtcatcattttgacaaattctgtcaacacacagcagtgctcacttggcactgtgatgggtcagctctgaacaga
cacacgcccctgagccctgcaaggtaaccagctgtataacaagaactccttccaccctgtgtctggaaacaagaagccca
ttagaccgggaagatcccgatggctattcctaagtgtgctggatggagtgccaggcccaactggcctgacctgtaagcctt
tcttccagctttgtgttctcctcccttgaagactccatttctgagttttgtgtgtgttttactttctagttgtgtctca
tcttaattttctctctctctctctgcttgaactgaaggttcgcttgggtgtggagagacaggcccccagcagagcagct
tcccagagacactcctccgatccagggtctcccagcagccgggcaaggcagggtctgctcttctgtctcagctcacaagca
tgcagggtcactcgtgcaagctgtctgtgtgttggggactgtccttaagccctgcacagccctgtcctcctgcccctc
tggaaattccaccctggctccacatttctgcacaaaatgagctgggttctgtgagcattggcccgccctgactcgcttagt
gggcccgttaagtggtttccacttcaaccttgcacctaataccagggtccacacacaggatggacattcattgagccgtgaa
tttccagtaataaataccacagatgcttccagacactgccccttctgcatacactcactccagccacttccagggcaaca
ggtaacagagacccagtcacaggaggggcagctgtggggggcaggactgcagctctccaaagcccatgcacaaaacccagcagc
gccttggcaggacaaggaggtgacattcagatgtggaggacaaggcatgaccattcctgggtcatgggggcccagcgt

FIG. 7

27/157

ggaactcagccttgaggcttggccagacttaacaccgctgtataaacaccaggaccttcttaggtagagtaaatggaacccaac
ctcaactgctcttagacagctgattatgctctctggagctgcgacaacaaatataccacaacctcagtgcttgaaaccaaga
gaatcagcttctcacagttctggagacagaagctcgaatgcagggtgtgcaggagctgagctctggagactcagg
ggaactcgtgctacctctcactcagcttcaggtggctcgtgacattctctggctctggtgctgcataccccaactccctatc
tctgctctccccctggctcttctgctcaaaatgctctgtgtttagtttccctgtagacacctctgcatcactcctaagatg
cagaggtgcgacatacaggtgttgagagccacttagataatccaggataagctcctctcaagatctgtaacttggtcgtg
gtgcagctgctcacactgtaatccccagcacttgggaaggccagggcgagggttcaacttgaggctcaggattggagaac
agctcggtggaacatcagggaagacctgtctctactaaatagccaggtgtgggacacaacctgtggtccagactctc
aggaggctgaggtgaggagattgctctgagcctgggagtttgagctgcagtggtgctatgactgcaccactgaattccagc
tbtgggtgacagagtgagactgtctcaaaaaaataaaacataaaacataaataatcaactcactctgccaacagaagt
aataactcttttgctacataaaggtaatatttacaggatccagggggttaggatgtggacatatcttgggaacctgac
agcctgaagacaactcataaatttcaaataggtctgtccttttattcttccagctcttgggaagcatatgctctata
ttttcaatccacaattctattttatttgaggatctcactctgtggtttttatttttatgagacagggtctcaactc
gtcacaggctggaatacactagacacatactggctcactgcaaccaactctcgggtcgaagtgtactcctcagctcagcc
tccctgagtgtcgtgaactagacacacatcaccatgctggctgagttactttttaaattttctagagacaggctcta
tgttgccagggtctggtctcaaaactcctggcccactgcaaacctcccgctctcgccctcccaatgtgctgggattataggag
taagcgcccttaccocagctccagcttttatgtgttttggtttggtttggttgagacagagtcagctctgacacaggc
tggagtgcagtgccagcatctcggtcagcaacctgctgctcggggtcacaaggtatctcctgctcagcctccga
gtagctcggtattataggcatgcacccaacctcgtgcttcaatttggatttatagagacagcggtttccacatgtgtg
gtgaacacaagaattttgagagctgggtttaaatacagcagcgtctcagaatgcaactactgtgtctcatcaagggaaa
ctgaactctgttttcttaaaacttatcaccagctggctcatcattttgacaaattctgtcaacaacagcctgctcattc
ctggttatctgtattggttcacatctgaagacacacgcctctcagtcagccctcagggtactggtctgataaacacgcctgg
caagaattccctccacccctgtgtccggaaaacaagaagtgtatgtgactcaactcagctgaacctccaaagtgtcggg
attcacaggcttgagacactgcatcggtctctcccggttttttttttttcaactgcttatatttacttaattaaatac
agtcaaaaaatgagaatagttgaatacactttcatgtgaaggcgaatcatttagcgtacttaactctgcatttgggtcta
ccatcgctgctgtggttagcgggggcagggtgatgagccctgtctcaacacacaccccgctctccccagccacttaca
cgctctcattccccagcaggtgtgggggcttagaggattcccccttctctgtaagtggagaatgggctggactcggtct
cactgcaccaacaaactcctttttcttttgggaacctgtcctccactccatataaactcattgggtgagatcactca
ccagctgataccgaagaggcccttactaattggccccactcagggtctcgtcgtggagacatcaataactcgtgaagagaga
gtgctctctccttactgagctgtctaaaggagacactgcaaaactgggtgctggtggtgctatctaacctcagtgtaga
gcctgcctggctgcaggggaaaccaatgacacagaccagcaggggcaacatcattgaaccctggagacagctgtgctg
aagcccaactcgtgctcaggaacatgctcagccactctgatttgcctgttcagtgcaagagatccttggcgggcagct
ggctcatgctgttaatccccagcatcttgggaaggccgagcaggcgagctcaactcaggtcaggagttcaagaccagctga
ccaactcgtgtgaacccctttctactaaaaatacaaaattagccgggcatgtgtggagatctcctgtatccagctgt
tagggaggctgaagcgtgggaatactctgtaaaactggaggtggaagtgagtgaggtcagtgagccagatgtgcaacctcga
ggcttagctaaacaggcgaggctctgactcaaaaaagagagctcctgactgaagctgtagaggtcgtgtggaacagctgcagc
aagcagtggtttgtgagctgattgctgagctcggaagtccacaactgatttcagagctgggtggctcttttccaactcaggaa
agctcagggctcgtgagttttggtgtggtctgtactaacactgtgtgtttaggtcaggtatctagaagccaaagcccaaga
caggagttctcaaatgcgctgatttactgaggggaatgccttcgggggaaactgcgcagtgaggagagctggagaggcggg
cagggggggagctgacagcaggtgggtcctgctggagttatgactcagcttgaccaccaaggggtcttgagacgga
acagcacactgcatgtgctcttgaggcagtcactggctcagttgctctgggacagagataaaagtggacagggcattct
tgggagtgacaactcctagagaaggggaacagctcagctgtgttaccagccacctttcaggggggtgggggattgctg
gcactggcccgagagaggtcttgagagagagccccacaacctcaactcaactcctgcgaacctcaagcagataga
aggagccgctagagacacaggaaccttctcctacagaaggtctccaagcttggaggaagaaggagggcgaggagggc
gggaggtcgttttaccocagggggttcagggtcgtgagtggaatgactcatgctcagctgactctgctgggtgacagat
gagacacctgtctcaaaaaacaaacagaaaaaaataatgctcttagcatctgctgggtccaacttgggcagctgtgtgtg
gagctggtccagctgcaagggaaggtggggaatcagtttcaattctctgagtagctagtgaggaagaagagacgaagga
atgaaattccctgcataaggaggtgagagctgattctatcagctctataagtgaggttagtggtggggccagacaggtt
tgaggccactggaggtgacacagcggcagagggtataccagcttgtagacacactgggtctcagggctccaggaacagaca
cacataaagacaaggtcagctctcagcacacagattataaagacactgtatgttatgctacagctgcaggaaaaacgggg
gaagatggccacagggccaattgccaggcaggggcacagggtggggccctcagcagcagtgagggaacaggaaagactct
gccagcatggtcagttgctgacaggtgggtcacaagtggtgggttaattgccagttgggaggtgagctgttagacgggtctt
gtccctggaattctctagggggttatatgcttgggaggtgcatacaattcctagaattctgggattctggatgggtctct

FIG. 7

28/157

aacctgtctctaggggatactgggtctccagagaggctctcagctacataaaagttagacctctctgcagcccccaaccttgttgggaga
ccccctgggtctccaaggtctgtgaataaaagggactacaaacaaagtgcctctgcagcaggagagcggggctttacagtc
agcccccagcctacatcccttgccccacagggaacggcacacagggtgtccaaaacctacaaagggcagcggagtgaaagcac
caggggccggcttggggggttaagcagagccctctcaaggggccagtagggcacagagcggcggccaggcaggcaggctggagtga
gtagctgcctgtgctctacagagggctgcgtctgcataacagagcaggggccgggagagcaggagcaaggatagagagcagc
ggagcgtctctccaggggccctgcctctgtggaggagcaaccttggggggtgggagccaaagtgcagacttgaggcmetaagcctgtc
ctgcctctctgcctgcaccccccacccaggcctctcatgctccacccagcccccacagggaacgggaggtgtgaaggtg
acagtggggagggtgcagctcacctccctccctgcacctactcagcaggaggtctgtgccgggtctccctgggggttg
cagctgcctccagcaggactggggggccacccctgtggcaaggctgaatgggacacatctaggagggagggggtccagggtgc
tagctagctgggaaacctagcacaggcctcagcacagggggacaaagtggagagacccctctgcmetaagctggctcct
ggacagaaaggtgttgttgagggaataacaaaaataaaggcacagctgtgcagagaagtgactcaactgtacagctctccaat
ctcccccttaaaatggctccacagcaggcggtctgggagggggcagtaagcgcgcgcgcccaactccccctcactcccc
cggggccctccacaaagcagcgttctctgttaagagcccccggaaggtctcagctctcagccgggttccctgtgggtgc
cgccagtcgcacaggagtcgcagccgcgcgcgcgcctcagctgtccacagtgtatccggggtaacggccgggtgcctcggggggccct
gaggggcgaagtgatcggggcccccacggcgaggagatccccctggagaaacccccgcctcgcccgctaccccggcgccgcta
ccccagaggcccgctaggcgcgcaggcagcggggcgccctggtccagcaggcagccgtcgggggcgctggggcgctggggt
gcacgagagcctgaaagggccaggagggcgaggcctgcgcggggggaggcggtgagccgggagagggcccgctccccacct
cgccacggggccggcgccgcgcgcgcgcgcacctgcagacccggctgcgctgcgagaagtttaagggctcaaggcaggaggg
cgccgcgcggcgccggggccgcgaaagggcgaggagggcgcaaggcttgaggcgccgcgtcgccggccagcggggccgcgcgc
acggggcagaagagcagtcgtgctgggtgagcggggcgccgcgcggcggaagacggcgccagggcccgccggcgccggcg
cggggacagccccctgcgcagccggccagggcactgcgcacatactcggagcgcaaacgcctggcacggctgcgcgcctgcgc
cgtgcaggcttcggcgccgtccacacatcggttccctccagccagctgcgcgcacctgggaagggcggcggggaggtgc
aggtctcagtcggcccgcatgcacactccacagccctccacggccctggaaggtggggagacatgaaacccagga
aagcatagagctgtggccaaaggcgggggccagtcggcactgatcggtctccgtctgggtagggcagcgggggcctct
ggagggctctgcacacgctcagggttggaaggaggagctggaggtgcaggcaggccttccccaaagggctgaagccctcc
ttcccaagcatcgccacacttccccaaacggccacacacgtccaaagacgggtccaggggctcctcctacactcaatggc
actcactgcagggggaaacatagagcttctggggatctgggccaatggagacttaaaacccagagatggcatgaagaca
gtggtcgtctcagctcttaatgcaggctgtgacgcatacctttcctcctcccccctcaactcgcgcttatctc
tccactaggagcctttgggttgagctattgatttgtgttctcctctccactcctcctcctcctcctcctcctc
cctcctgggtgtctcctcttcagcgccttctgatgctccttgagctcaggtctctcctgccactctgcactggtct
cctccagccctccacctctcatcagccaggggagagcggatttgtaagaaagctctggccactgaagaacccactgact
ctgggtgcagggttggttagagaaagtcagtcagcacacccgcggccttgtctcagtggtcagcagcccaagg
ttcactcagtggttttagtgacagggcccatgtagtgtagacacctgtggagacagagaaggtctcggcacccaggt
acacacaggaggttggaagcagaggggtgagctctgccacatgccacaggtgtgtctaaatagagcagggaataact
ggaagtgtgggaactggcagcccttccggagtttgagtcacaaaaacagggctaatcgaatccccccagcctccaca
gttctcagctctggagcatctcatcagactagaagcctttctggtgccagccctgagtggaactgacagcctggcccc
tcatccatctcctctgtgttgttgccaggggctacgggtgagtcccccccaggctgggagcagggtgcctcctc
ggcatccggctgtgctcttcaggctgtggattggacgcgctggtagagttctcctgtgtacttcagcactgcctct
gcaggagaggtacacatagctcagggcttgaaatggctcagctgagtgaggggatagatggcagcggggtagtccctcc
tctcctccctcccaaaaggccacacttgtcttctctcagcagcttgccaggagagactcagctctcagctgcacaca
ccagcactgtatgttgctaaaggtttgaaacccacctctgctgtgttacctgtcacccttgcccagaacaacctac
ccagcttccccctccacagcagctcatctcagctcgaaagtctccccacttaaacctttaccagattccacttccac
actacactcggtaagtgacacctccctccgcgacatccatgcaggcagcagcagtcagaagagggtggaacaacagatg
aacaacacccagatctcaaccaggaatctgttgtgatgggtgagagggtcggggtgaggggaggtgagagaagttgggag
atggcagacaataggcaaaagtctcctgggcacgggtgcgaaggctggaactcgggaaggggagtggtgtgtgtgac
gggggtgggaggtgaggtgctcagcagcatatggggctaatgtgccaggggtccccagccgtgcgcacaggcctctcact
ggcacacctctcctgtgttcagctttctggaaatctctgtccctccaccttgtcactgaggagcagggacagagaggc
cagggttcggggctcagggggacagagggggcgtccacagtaaacagcagcagcagcagcagcggccaggcagcgga
catgctggacactcaacagagaaggttggtggcagagctcagctctgataggcgccccctctcactgctctggagca
ggagcacagtgaaaacccagctgtctgcgtgatccccctgggaggagctttgtgcagcaactccccgcctcataa
gtctccccccagggtgcacgctcaggttaacctgggagcagcggatcgccgcctgaatgtgctagacgggacccca
cacactagcctctgccacgcacacccagcaacacagcaaacacccaggagagactggtctcctcagctcgaataag
aacgtatgaacagcagcttttcttctctttttataatacttatcttaaaacttaaaaaaattttttgtaagagt

FIG. 7

29/157

[illegible]

FIG. 7

30/157

caactgcagtggtgaactcaggctcaagtgatcctcctgcttcagccctcctgagcagctaggcctacaggtacacaccacc
acgcccagctaatctttaaattttctctgtagatgcgggtgcttgcctgctgtgcccagcctgggtgtggaacctcctggcctc
tggtgatcctcctacccagcttcccaaatgctgggatataaggcatgagccaccacgcctatagccaacatgctcttc
ttttgactctcactttggatctcttttcttaaatgggtccctctgtccccccgacacacagaaatgggggagaggctgtc
agatctctgagctccagaacctcagggtgtagcactgggatgggggtgggggctcaggaaacacctagggggagaagacagg
gtgggaagaaaacaggaaggaaggtcccaaaaattatgtttgtttgcagaggccagccaggctccaggggaggtgtggaactc
agtgcgaacctatgggccccaggaccactagcttctggccagcagctcatgcctccacagagctgggtccgtggaaattgc
atgtaggagacacaccagactccaggacagagccctttgggatggccagcactaccagccctccactggtgagggaag
tcaggggctgtgtgacctttgcttctgggactgatgggttatctgagctggagagtggtcccagcagtgctctccagccct
caggaaacttctagtggtgctgggttctctggagtggtgggtcgaaagctccactcgggggaagaaacttccagctgcct
gcaggtgctggaggtcccggtgatctcagctctgcccctgcagttcaagttcctggagtggtgctcagtgccacctgt
cctttaaactgttctcatcttaggagctacctctcaccagaggcaggatcttgacctggagcttgactgctgagaatgag
gaggatagtgtgtccctaaaggactggggccccagggtgcaagctgtgtggcagagagccctcctcactcagtgaggac
cagtgatccaggaagcaacagcttctcctccccagccagggtctccagcatcctggctcccatgataaccaagag
gctataaactcatttccataataacctgagccacagaaacctgattagggggcagcaactcagggggtgggagaggtggga
gggtggcgatgagaggggggaggtcttgaactcagggtccctgcctaccttgggggtcagggcgagactgctggcagggtct
tctcaggggtggctgctgggctcctagagagtctcagggtctgggagaaaatggtgagggttaaatgttgtgaataggtc
cagggagacccctggggctggggaggggcagataggggactcaaggtgactgggtgctgcccactcagccaggggcccagggtgctccagtggt
atgagcccttcccttctcccttccctctccactccccctgggtgctcactcaccaggggcccagggtgctccagtggt
gtggggcccaactccatgggggtgaacgcgcgcagggggtgggtccctgtgtgggcaactcttggggctgagcaactgtag
aagatccagggaggttgccacagtgatccttgagtggttatgtcctccccgagcatgggtgtccagccaggaggttctg
cgtttactgagtttcttggggcaccctactgctccaagtcacccctctcagctcccttctctgctcccttctcaggggagcc
ttgggatccaggtccaagtgagctctatgcctcgggtggcactccctctctctagctcctaacttctcctcaggtct
gacacacccagcagcctggcactctccagatgctggcatcgctcagcttccaaagaaaccttggatgtccgccccttcgg
cagctatgtctgctccttgcctcctgggtgcccctgctgccccttgatgatccaaagcaactcttgactgtccccatcca
tccccaaagggccttgccttctctgtgtagtctcttcaaaaacttccctcagccctggagactcccgcttggggatgaga
agcagccgcacccctctcagcgccccctcogtgctgacacgcagggctctggccacacttgcctcctctgcccacagcc
tcaatacaactcgcttcttgcagggtcctcttagctgccaacccggggccagggtgggtgcccctgcccctgctgtatgccc
ctctgcccccatctctggcccaaatcatgccatctcccttggcttgcgggagcactcccaagacagggtatgtcagac
atgggacacagagtgcctgcccctgctcaggccctgggtgcagggtgagtcctaggacagccatgcttagtatgtgact
ccccactccgcccacccagggtcacagagaactgggttaaggcagggccctggcagcaggggacgcagcaccgcagctg
accagtggtatggagtgaagagatgtgctgggcccagcatcttgggaactcgaagggggtgacagaggtgatctgtgcaga
ggaagtggcaagggccgaaaactggtagagacaggtgggacagggctccgggggacagcatggtagagggactgcaatc
tgagccaggggaaaaacagggcgaagtcaaggggtgagggcagccagctgggtgggagaagcaggagagtggaacaagaggagct
gtactgggaggtagaggggcatgcttgcgggtgctgggtgggtggggcagggatccacccccctcgcttgaactgggagggcac
tggaacctcctgttcaagactacttcttccactggcctctggggctgctctctgcacctgggggcaaggctgagggcctg
ccccagctcccaacagccccagcagactctgaggagggggaacccgagggataggctcaggaagcaggcgctcggagccta
ccccctgcacgcagggtcccttctgcagccccagctgcatcgctgcagatgggtcctgggagtcggtagcaaacaccagg
ccaggccggggccctggggacagagggcagcaggacgctgaggagcagtgggccagcaggaaggtgtcatggctgcccgggatt
gggggaaggggcgctgagctcctgagcaggtgcacccacccagctcctgccacaccccccactggcatacttccagcc
tgcaagggccactgtccatgctgccaccaaagcctggctgtgcagagaaggggtggagccagcctggagcagcagtgggca
gttatgggtgtgatttcaaaacctgcagagaagagaagaggaggtcagctaggataggaaaccccaggtgcaccccact
cctggggtcctcaccctgctggctgggacagctcaggcagcgctgaagctcctggaagcatcttctctgcagcgct
gctctggcacccctcaggggcagcagcagcgagtgagggggacactggctgcgggatctcaaacggcagggagggtgcg
gggtgggcagagctcaacttctaggtccacactgggtgcttgcctctacacagccagggccagggctcagctgggca
ctggggagtgccaaggctccccgacaagtcaactcctggccatctaggtgaacggtagagtcactggcaccatgtgcgg
tgagagcaggtgggttccacccctggcactgcagctccacacaggggcatccctggggaggaagttagagggggtcga
acagctgcagtcagctccccccttccccaaacccactccatagcttctgctccctccactcagctccactccctacccctgc
acagggcaggaagtggcctcgctgtcctggcccgagttccatgagcatctcccgagaggttccacccctggaaacagg
cctcgggagctgggtgggagctcaggaagcatggggcaggtcggggcagctcggagaggggtgcgctcagcgggcct
cagtttccctgcccatcttccccacagttagaaaactggccccaacagaggaatggggggcaggggtgcaaggggtgctcgtg
tccctctcaccagggccccagagctgctggcactgctcctcagcgtgggacatcgccatccacagcagtagccactgccc
ctggcaggggtgagcgctcagtaggttaaacgtctggggacagtgaggaggaggtgcccgtgcaaaaactcagggaaggtc

FIG. 7

31/157

acagtcacccatggcctggcgccacagcgctccagccgggttcagctcgcgcaggtgacgggtgggtggggaaggcagagag
aggccacgctgagctgagaggtccatgcccagagagcgcggtccggagctggggggagccagggcctacccaggccagcaccca
accggggaaacctgaggggccacaaatataaaggcccaacaggccgggtcccaagctccccgaaacctcaccctgaacct
ccatgccctcaccaggcagcgacgcagcgagctccccgtggggcgactggggccccggggcgagcgagcagttgtgagcaa
agcagcagaggtcgcgggactcctgggaccagaaaggcaagaaggcccgaggtgagggcgagcgccccagacctgagcg
gagagggcaagtgggggggcgggcgagcgagcttaacctggccaggggcgagtcacactcctcgccgcttccacgaagc
cggtccccgcagagcgccggcgccacccggaggtccgggggtccggggcattggagaggcagcgccgcccccttgcgggaag
aaggcgcgagctggcgcggtgcaggcgctgaacacgcgcggaacacgggtgctaccggcagcgggagggcattggggc
atggagggacagtcacccaaacccccgcgttctctgatccccacccctgggcttgggtacagccgcagacgcgcagagc
ccagagaggggaagttaacccgcgcaagtcacacaaacaggcggaacaggggacgatggggccccaaatagtgagcagcccg
ggacccaaggtggaatcgcgacccgacgggtgctcctccgggtgtaggagtaacctcgccagggttactcggaaaataatct
tcataccgttgagaatccactttgctgagcttcttcccttaagcctcataaaacccctgaagcggacactatgatca
ttatccccatttaccagaagaggaaactgagggacgcacaaagaacacgcagcgagggaagtccccaggactagccgcccc
ggcgagccccgcacccccacccgcgtacccgggtggcgcgagccatgacgcagcctccggactcggcgcgagcctccacg
cagcagccgtcggggtcgtggctgaggcgagggtctggccgatctcatgggcatgtgtggctcgggcgccgatggggag
ctccagttggtcctggggggcggtgggagggcggtcactgcggcgctagagcctcctgtctctccccgcgcccgccgcg
gggggtcaccgtgctcagccctcccgagctctcggcgcgccacatgcctcgagggcgccagggccactgtggcgccct
ggaaggcgcgccctggggcgagcgcgcggtgaacaggcgggggcgggaggtgagggcgccccacccggcgccgcg
tcgggtcagagggcaccacgtgagcagctgcggggagtcgtggggcgctgcgccccacagccccggcgccactgacgg
aaggccccagagcggcgctggcgttggcgtcctgcgtgacgcggctcggtccgcgtcggtccacactcctccagccggtgacgc
cacctgaatgtccagagtcctgagaagctgaggggcagggcggggtgaagccgggacagggcgccccatcgccggctgg
tcccttctggggcgcccttctcttccccaaacccacagcacctgctgtcctgcggcgccaccccaatcacccgtct
tctccccgcgcccccacccgttggccacgttggcgacttccaggagacgctgttgggtgggttcaagtttctgggtg
cgagtcagaactgggaaggcagaaaatcccggtggcttgaggggctgagctggccccatccctgaccccccaacccctg
gggtctctctacacagggtgtggtctgccaactgtacagttccaggtacttcggggtcctgcgctctctcgccctg
ctgcggaggtgcaaatggggaccctgagtggaagctgctggggttgagccctgaccccccaacccagctccccagaaggaa
gtttaactgtttctggaactgtttcttcagactcaataaaaaatctgggactcgaggccctgtgaattcctgtctct
tctgatttggagggtctatagatacagcatccccactccatccgatcgatgcccctgacccctgctctggggaccacagg
aaggctggtccatgccgcttgttccccaggatccctgtggccacaggttcccttccagggtgagcagctgctccatccgaa
agatctcgtgggttgagaagtccttgaggccccgggtggccagggaacgcagataatagctggcttctgtcgtgaggtg
atcaggcctcactagggtgcagaggggttaggagcgggtgtgaggagctcttccccatccaggccacgctcctctccca
gagctcacctcatccacagcaggtgcagaggactaccaggagtcggggaagcccttactcgcccttggtagtggcaa
tgatcctaggggaggaaggggccacccccaaatctcagccagggtggagcaagaggggcaagaggggaggtgtggtaggg
gctgggtccacccgcccccttaggaatgcaaggaggagtaggggtaggaatggtgggggggtacctctggcggtgcactccc
agagcccatggaagcatctcacgtgtggttggggggcagcaccaactggctgccatctgggcccgtatgggttctatg
tatcctggggccagcagcctgctgagaggggtgttacagggaacactgaattcagcttccctcgccctccacaggatg
tctccagccttccctcctaaatgctaagtgagcagcttctatgagtgagacactcacagttgcttatagggaagggaacag
gagcaatggtgacttgctcagatcagaaaactcttggggctagaggaggagccttggtgatggcttagttgtgggagatg
tgaattgggaagcaccaggaggagcggggggaggtgggaatagggggaaggtttggtggttccccagggaacctgcag
caggcagaggatccacagctcgggagggaggagtcaggagacactgccgaagaatgggaactgggttggggaaatg
cggtgacctccccagttccccctgctcctgcctccttcttgggtcagctcgtgagcctcttgcgcccaacctgcctca
ccttgaaagggtaggagcgtgggaaaaatatggttgggcaactggaaactagatttggctcctgtccatccgatccca
cgagggaataaaaaactgctgaagtgtcagaacagcgctcggtgcaagcactcagtaggtgatatacatgataaac
ataatcgtgggtggttttaagagcctgcctccagccgggaacccccaccagccacccctcctgagcagcag
gtctgcctgtccctccccagctcaaaatcttgggtcctggatggtccaggacactgagacccaaaaatggcctctctgtga
tctggcctcctggagctccaaaatctatccccctgctccccctcctgtagatggagacctccaggcaggtcagac
aaactgctcccccaaggttagcactgcatctttctttttctttctttctttctttctttctttctttctttctttct
ctctttttttgagacggaggtctcactcttggccaggcgagggtggtgagtggtgtgatcttgggtcaccacaaacctct

FIG. 7

32/157

acctccgggggttcaagcgatctcccgtctcagcccccctaagtagctgggattacaggcgccgcacacagccttggtctaa
tcttttgtatttttagtagagacagggggtttcacatatttggcaagcctgggtctcaaaactcctgacctcaggttgacgcgc
gcctgcctccctccaaagccacgcgcatcttggtccctgccattcccttagcctggggtgcgcctcatcttttccctctat
ggattctcttagactcagcatattcttgaattctccataggttggtgctcactcatcgccagcaggaggtacaacagcgc
ctctctgggttgggaggcgagggtgcccacacagcgccgttgacagctcagacggttccctctcagctccaggggagccaa
ccacagcctgcgtgactctccctgtcatctgcacccctctctggggtctcttgcccatccagccacccgcacagatctgt
gtcagctcctgcgcccccacactgatccctctccctccagcctatcccacagcctggacactcactgggtctctctccagctca
agcaggagctctctgcctctcagctccaggggccaccagccccattctgtgcttcgagacgtgggcagaagaagagtggtga
gctgagatggtggtcctccagcctctcctgcctccagggagtagtggtgctcttggaacgcgtctggagcggcctggaggaagtcttggg
gagaaagctgggtcggggagtcagcaggaaacccccactattatggaggcgctggaggagtggaacatacaaaagattct
cccaactcaggatgctttgcagtcctgacaacacagccatttggaggggcgcttggcacagcctgccaggccacacagcatct
tccatagagacagaggctccagaaagggttgccctctccctggcccgcctctcttcttactgcccagaaggggcatcaaa
acagggggaagacagagggtgtgaggacattatgggggctacgggttgacatggtcagggctctcagagcctctctgc
tacctaggcttggttcccaaatgagctgtgctcatcttctatgaattcaaatctgaacctccactcccaattcttggc
aaactgcctcctctccaaagtcttctgggctccagcagcccccctctccggctccgcacactgcttctactggacc
cacgaagtacaatggacgcaattctccagccaaagagagcaactggctctcagctaggtgtcaggaggtcgttctggacgg
ccagcctctctctctcccccacccctcttggtgctcccccacttggggaaccccccaactcccccttgctcacagctc
tgcttgggggccagagtggagcgggccaggagcctctgggaggtctgcgccttggaattctttaaccccgaatc
cctttctccaaaatgcactggcctggggcctctcgaagggtctcagagttcttgaggggagtgctctctctgcgaagt
ggggagcagatggtctctgctccctggccacagcgcccacacaggcctcagcatgagctcagagctgaggttggaatgccat
tgcttttttggggaaaaggtctgcacgggaaaaggccatactcagaggtcctgttctctgcagccctgctactcttta
ctcttgcctcctggtacagctccctctgatatataccctcatcttgaaattgtgagttgtctctgcctcttggaggggga
tacctagctctactcttcttctgaaccttggcaggtctctgggggcaggggccacaggtgggggaagcagagc
ccctttggggctctctctcttggtcacagccagccagacagacagaggagggccacagggagagtgaccacctgctgtgt
ccagcctccccctcctgggagtggggaaggccaatctcaaaactcaggccagctgcgtgcttgaccagtgaattgggggc
ttatgggctcaggggatccagtgaggggcctgggttgggagctgctgggtctctgggggctctcagcctctaatggc
gtctccctcctctctctctcttctgtggaatttggaacagttaggttgaaattccaaaacaagagggtctctaggaggggga
ggggttgtaacctctgtttaaaactggtcagacacttagtggtctcaggtccagcctcaagagctgtgtgaccttga
caatttccccaaagctctctgggcttcggttgcctccttgtaaaatgagcatatacagggtcactgcctcttagtttgca
gcacagatatttgccacaaaagatggggcatggttattctggaagcgtgtgaagagcgggattgggaagagagctgggg
cagagcgtcctgcagaagaagcacatggggtggtcttatctctgggggacatcaggagagtgaccactgccccccata
ccagaagtgaattccacaggagccagtgaggctgaaggttcaggcctctcgtggcaggggccctgagagggacagcagctgtg
tccacagggtcacatgcttctgtgcaacttgaaaaagggttctcttttgggtgcttttttttttttttttttttttagag
ctcctgaaaaggtctcaggaccccaaaacttgaccatttcttgcctggtgggggtgggggtggccagatctcaaggg
agggaaggaagaggaggtgggggtgggaagagctgaaatgactctcatgtgtgcggctccagagatccagatgtccaa
accccaagtgcctctctctgccacttgagggggcaggggaggcaggggcctataggagtgtgacttggtggtctctgggga
ccccagcaaaactgaagaagctgtaattgtagggagagacaagggtctggaggttcaggggcccttgaggggcggggaga
catggcatcagcggctcctcagagctgacggtgcgcacgggtgtctcatcaggacccagtgcggggtgactgctgc
cagggaattgtcctggagtagaagacagacaggttgagggggacatcaggagaaacacaggggcttggaacaaagcagag
gaggggggttagaggaactccccaggagggagcctggagggccttttggggagagctcacctcaaacagcgtgggtcagc
ctggagaaggagggacgcccgtgggcatccttggtctcaggagagctatcaaggaggaggaaaagagaaggcttggaaggg
acagctcagctggggacagggagtcctcctgactcttgctggggggcagggtctgggtcgggcatccaaaggagaagggc
aaggccgcagctgcagagggaaggaagagcgcgcagcttagggattttagatggggcagagatgctcttaggtgt
agagatgtacaggaagagagacacttgtgcccccccatcatctgagaaaaaacacagcagatgttgccttgaggagttc
accttcccagagctcctcgggagctgtcctggtgtcagggttttggtactccttgccacagaggcccccctcatct
cttcaaggaggagggaagcgtctccggaaggagcttggtgtcctcctggccgggttgctcctaagggggctcaggaggaa
ccagagccaagatctactcagaggcgctgggaattgtcccttggaatgctccagcgcctccactggcccccaaccctc
tcacaggcccgccctgcaggagcgaggccacagacccagagctgcagcagcctctctccccgcttaccagtccca
gctccagaaacagacagcctccccctccacgctcctggctcagctcgtctgggtctgtggtgctgttggaagtga
ctcagctctgtctaggccaccccactccttttctccacactctctccagagatacaaacatacaagacactctc
tccaaagagcctcttgattggatgagtgaattgccactcaggcagatgaggggcagagagagctctgcaccttggaag
gaggtctagaggggccagctgcaggaggaggtctgagtggtctgggggagggggaaggaaggagggtctctcagccacaga
gcaactcaactgagagttagaacaacgaacttgtctgtctctagcctctaaaggtcttggggaagaggttaggttgggccgc

FIG. 7

33/157

gcacaggtgtggtgtgggtgcagtggtgtgtgggtgtgtgtccacatggccttgcgcgcacgtgctggccacggggcaacc
ctgaccccacatgagggagagaggggacagagctggagctggagctccggtgacccgggtgaatgggggtggaaccc
gagggagccagagctggtattgggacacatagacgcccctctccaggggtcccatcacctccctgaccccaggtatgggc
tcagaggggagggagcagtggaaccgctggggccctccctggggccagaacacagacagggccctgtacctgttgggtcc
ccacacagtgctgtggaagccacgcccagctctgcatagcacagccacgcccgcgatgcccctccctggttgcctccc
tgttcccggccaggcacttgtgtgcaggactggctaactcctcccacccgcttgcagaggttgttccagcccactctaa
catcttggtttggagggttaccgcgaggagacagctgcagctcttccagagcactgctaaacagacacactctctatctgg
agaggcccttctctatctcaacaaaggcaacataaaacaacatacacactgcccctgctgcccctgggagggagggac
gaggggtgagcaggggtggaggccacagctagtctctgcagcctgagagcaaaagcagggaactctgggggactcttgggcatg
ggggcttctagaggtatggagccccgctgagctctaagggtggaggagcaggagcgggtcacacggttgccctgcccgtg
gaagctggttgtgagagcgagaaaccaggacagaggggctacggctatgggtgtgggggtgggggtgggctgtccccga
gggggagggagccatgcccctctgcttctgcagcggagtgaggcagggtgagggaagtccgggcccggggccagcc
caagcactctgagcgtccctgggcagggtcccacggagaccccccaaggagtccccacgcccctgacctactggccgtat
gggtccggggccgtgagaccctccgcgcctgacccgagctctgagcagaaacccatccccgccaccacacccgcgctag
cctgcccctcaggggccacccccccgcgctcctcaacttgaagcaacccggcgccctggcactggccagagcagcagct
agtagcagcagcagcaacggggtcccccgagctctccggggccctccagcccactagctgtgagctcctcgccctctaggca
cgggctcgcaactccggctccgcggaggtggattgcggccgacccgtgcccgtgcagcctcaggccgcccgtctcgga
ccttcccgcgccacactcccacgcgcgcctcgctcccgcctcccccccccaaccccgctcggagcctggccagggt
gccccgaagggcgcgccatggggggagccgggtgcgcaactcccggacccgcgcgcctcggaggggtggaagtcggcgag
gagggaaatccgtgcggccctcggtatgacgggcccagccgtcctcccgtcggtctcagaggggctctactcctggaga
ggaggagagaaacgctgggaaggttcttggaggaacggcgcgctgggtgggagggcggtgggcaaaagggccgctctcgct
gctgaagtggccccaggagcgcatctcctggtgtctcctggggccgatctctgtcccctccttgctacccgctcctgccc
ccgaggggtgcccctggcgagggttgagtcgggtcatccactgcactgggtgcccccaaggataggaaggttcaggcaacc
gggtgcgcgctcttggggctctcatgtgcgtcaaaaggcgatgcagcagacggagacaacacttcttccctggcggtgg
ccagagggcgagaattgcataaaagctgcgaactcccaggcctgggagaccccttccggcctcagtaacatctggtttcatgt
tttaaaacttttggtttccactcgggtgcgaacttggatgagatgttaacttttttttttttttttttttttttttttttt
ctccctctgtcgccaggtggagtgccagggccgatcttgggtcactgcaacctccgactccctgggttcaagcgatctct
cctgctcagcctcccgagtgcctgggactacaggccgcgcctaccaccccagctaaacttttggatcttttagcagagac
gaggtttccaccattttggccaggatgggtctcaatcctctgatctcgtgatccaccgcccctgggctctcaagcgctggg
attacaggcatgagccaccgcccgcggcgagagatgttaacttttaagcaaatcttttttttttttttttttttttttttt
agagtttctctctgttaccagactggagtgcaatggcatgatctgggctcactgcaacctctgcccctccagatcaag
tgatcttctctgctcagcctcccgagtgcctggcattacagggaattcgccacccacgcccgtggtaattttgtatttttagt
agagatgggggttctccatgttggtcaggctgggtctcgaactccgacactcaggtgatctgcgcgcccctggcctcccaaa
gcgctggaattacaggcgctgagacaccgcaaccagccactactttaagtaaatctatttggtttttgagaatttggaaatgta
gtaatttgggttagtgaaagttcgagcagtgagagaaacactacattcacatatctcaaaatcaaaaagtacagaaagcata
gggaaaagtctccgtgctctagccctcctcaccaacaggaaaccaaatatgattagtttcttcttcataggcttttagatta
ttttttcacactcaagacaatacagacatattttttctcttattaaagttttctgcaactttgattttcttttttttttt
tggtcgcttaatacaccttagatatcagtggttttagagggtccttgttgttcttatgattattatttagagacagggtc
tcaactctgtcaccacagctagaggacagtgccgtgatcatgcccctattgcccagcttgaactcctgggctcaaggtatcct
cccacctcagcctcctgagtagctggaactacaggcaacggcaccagggccagctaaaaattttaattttctgtagac
agggggtcacttgtttgttccagcctgggtctcaaaactccggctctggccaggcgagtgctcatgcccgtgtaatccca
gcaacttggggggcgagggcgggcagatcactggaggtcaggagttcaagacacagctcggccacacatggtgaaaccccat
ctctactaaaaatacaaaaattagccgggactgggtgagcgccctgagttccagactctgggagctcaggcgaggaa
aatcgcttgaactcagaaggtggaggttgcagcgagccgagatcatgcccattgcaactccagcctggggcaacaagagcaaa
actccgctcaaaaaataaaaaataaaaaagaaactcctgatcttaagtgtcctcctgctcagctcctcctcaactcgt
ctggaaattacaggagtgagtcaccacagctgtccagctacgagattattactattattactacttttggattttcaaatc
aactcttataaggtataattacacacacataaaatgcactattttaagtgccagtagaagtaggttccgataagtgat
ataactacataagcatcactataatgcagacacattcctcactcacgaaagagccctgtgcccctccagccaaacttc
ccactcccaacccacagcagccactgactctgttgttctctgtctctatagataaggttttgccgtctcgaattttcatata
aattggaatcagtcggccacactctctctgtgtctggcttccctcctcttctccgactcttttagactatttcaact
tttgcatatcaatagtttgtttcctctgtattgtctgaatagtggtcggtgggtttgagggaacacaggtttctctactcacc
agtgacacataggggttatttccagttaggggctcttaaatggaaactataattgacagagagagagaggaagaa
gagggagagagatatatttattagcaattggctcagctgattatggaggccaaaagtctccgaactctgccactcgcaag

FIG. 7

34/157

ctggagaaacaggagaaagccagtggtgtgatttcagttttagttcaaggccctgagaaccaggagcacagtagtgagggtgg
ctcgagctcagaacaaagtctggggacaggaaagcagagcagcacccagagcagccctcagcgacacctcttcagtaaa
caaggctgaacacagaggggctggcttcagtggtgtagtcaggtagacagaaggcagctcgaggagctactctggcgctctt
gcttactggtattcttacctcgaactggccaaactcctacttaaacctgcaggccatggctttaatgtcctgcoacttcagag
gctgtcccttaccocaaagccaggttagcatccctgactgacacttctcctgcaacacgcttcagaaggccctgtagtc
gtccactccctgtctctcctcccaagctcctgagctccatgtggtctgggaatatgtgtgtgtgctcactccttagcaca
gtcagtgctaaataactgactgttagaggggacacagtcgaaaagccacatggggatcagagtcactccttacacagtgaca
cctcccaaacccagatgagctgtgtccaagtgcaggtcagaggaatttctgcccgaagtctctgagaagggtttatttta
cattttgaggttgcaggggaggagatgaggccatcaaaccaagctgagggaaggggatcctaggatgcacagcagcagct
ccggggggcgctgacagcacctgggaagatggctctcctcactggcttgttggcgctcaccctccagagggggcatcaggaa
atgtcctgggaaccaggcaaacccagtgagcattaacccttagaagtgtctggcatgggtgaacaccccatctgtaaaca
cgactctcccaaggagtgacgcgaacacaggatgtctgaggaggcactcgactccagctcagagatcgccagggtg
gcacctggtgcagcagagctgtagcttgggtgcccagaaaaaggtcactgtgtgtgaatgggggccccaaagccaacgctt
catcctgcagcagcctggtgcoacttagaggggaaacttttctccttggcaaggtgggtggaatttcagggttccatgggca
agggtattttagctttaatagatatgtcaaacagttttccaaagtcactgtacacactctgtgattctacttatgtaaa
gtttaaaaacaggcaaatcaaatctatggtgttgcagactcaagcagtagttacacttctgggggctcgaactgggtacag
agtgtaagggggggactgtaggatggctctatttcttgatctgggtgtgttcgcttttggaaaagtcccttgagtgcattta
taattgtgtgaactttctgtatgttacactttaattgaatgtacaaaagtctcaggaggctcagaccactggaagcgg
acacaactaaacccctctgagagcctccaatccaagatggacatatgccccttggaaagtatgcagaagcaggtgaagact
cctaagccggtatattcccaaatccccccagtagcgcagcttcagcagctgcttatgggtcctccctacacccctctctcc
ccagacagcccccaaacctctggctgcatttgcattgctctcctgtccactctggatctagtcactgtctccac
cctcccccactgtcagcaatgtgacaagacaaaacgcttagttcacgtgcccactactgcgtgccatgcacggggctggt
catttgggtggcaaatgtgagcaacacacgaagcctcaggagcagaaaagggaacacaaactcactcagcgttaaggtaat
tgtgtataaattcactgtgaacttcagcccccctggcccccctcctacagatgggtgtctaaagaaataaacccactaacactgt
actcctctgtctcagccagctgtcttgggttgcagaagagagactcactccagttgcgtctcagtaggaggtttatttgg
gaggacattctggacgggtccagcagaagactctggcaatggagcatgaaatgaaatgagctggaatgagggagggtgc
ggcctcggtacacaaagtccacccggcccccaactgctccacagcgtgttagctcctgtggcactcctccactctctc
tctgtttttcaaccccaaatcctagagcagagggtctctcgtcctcacacactctccacaggctgcagggggagtcact
agccactcaagagctctctcctgctgggtccctgtatgcaggggcagtcactccttcttggttaactcatctctttaa
cagcccgacacagggtctcagggaaaaggacaggagctcacacagtcactgaacagagatgcatctggagcaggaataagg
atgctgacaaactctgtgtctgcccctccactctgttaagtgcagtaaggtagtaccagctcctgggggtagggggggt
ggtggtgggaatgagggtcactctctcttccaggggccaaactctggcacaaggatgtctgcttagggcaactccaactcctg
gcatgctctcctcctaggaaaaaacccaaacactttttatttctcactgtcttagttgagagtgatcacctctccaggaa
gccacacccagaaatggccaggggcacaccctggccctcagtagatgggcaactgagacaaaacatggcctgctggtggca
ttcccaacaaatgtcaaaagtctcagggtgagctcacagtgggagtccttgggaagtccaattccaatgtcctctgga
ctcaaaaataaaaactacattccagctcctccttgcagcttgagtagtgcagtagccatgtgcaaatactccagccaatgggaagt
agtggaaaagtctctgtgcagcttctgagttgogtccaccatgggcatgggtgtgcttctacactccactctccttctct
ggctggagggtcatgtggacaaagcggggccatggtagcctcaaatcaaggctcattttagggtatagaaaaggaa
gatagaaggatcctgagcccaacatcataaacctcgtaacagccagactagcgtgggagagaaaaataaaatctctat
atgttgaagtcactgatttgggtctctgttagagcagggtgactgatactctgttgaatacagaatgctctggtgagcc
tctgagatgcagtagtgcactcaatgcaggtcagggtagcttctgaggagaaggaatgtgaaggccctgggagccgggtctcc
tttggggagaaaaatggagagaaaggaggagatccagagaaagagagagcagatcagctcagtgctctcagtgctcctcca
gggtcaacacccctcctgggcagactctccaccacccatttttgggggggcaagaggtctggggccaggtcgtcaaaaaag
agatgtcacagagctgtttctgtagaggcgggcaggactcaacactgctcattcactcactaggtcggagctacaca
gagctcggccacttctcctcggaggggcaggcaacaccactggtcagccagggggtgggcaacaggttgaggtctcaaca
tgtggtctgcatcaggatcaatgagctctccacagagaaaacccctggtcacagagagccctgggagccaggggcaggcccc
caggccccagccctcaggtgggggtccccaggagatgtggggaggtgtgcatgtgctggaagtgtgactggagccctggcc
taacaggtgaatcagctgcagaagggaacacacacaggggagagaaaactgacttctcactctgcccaggtgacggggcag
gcttgggagtggggaagaggcctcagaacaaaggtggggagtcagaaactcagtgctcttggaaatggaggggagctgctg
caagtcacagaagtctccctagtctcttaagccagccgcagctgacacagggatccctctgcccagcagacagag
gctctcctatctcctgaggggctcctgggttccctgcacactctcccaacactcccttgggtccaggcaccatcctctaa
atgcagcagctccagatgcagatgggaactgtgggttagctgggtatataatctgcccgtgtgtgtgaatggataacc
tgctgtggtctcttttgaaaagccactccacccaatctcgcccatgctctgcttaaaaaacagcctccctctgcaagg

FIG. 7

35/157

ccccggagtggactcagagagcctctccctccccacctcctgccccattcctctccgctccccacaggctggctccag
accacaaggagctgggaacctcatggagagtggtgtcttctctgtcttctccctagcccttgccttagcacagccagg
gcagagaggaggaaagaaaggaaactacagcaaggaggatctagggaacattctagaaggatctccaatttgagggggca
gtctgggaagtaggcagcctgtgaaacttaattaggaggtcctaaaggacctgggggaatttggggagaaagcagagtgggc
tgcttaccoccatctctggccactcccagctccagaccacctccaggtgtagcaggccccagggccaaacagcagg
agcaggaggccccacagcagctcagggccccagctgcttcttccaaggggaccacccaagatgacacctcttaacct
cgggactcagagcagccacagctgcaggccccagctgcttcttccaaggggaccacccaagatgacacctcttaacct
ttgctttatcgctgaagctctccacacagggtgcgccgagcaatcagtttagagtcgacaggcaaatcatgctgcagca
ggagggtatcacagggaaggagccttggggcctaaagagctgggcttttgtgcagagtgggcctgggagatggagccccctcc
ctccacctctgaccccaagtaggtggagggtagaggcagagccaggacatttagggcagaacagatgtattcccccttgg
tcgctggggcctcagcgctcgtatgctccacctcaggggcattgggggaaggccaggacatggatgtgtggctctgcaggagc
accctggggctgactggaggccaccagtcgactgccaagggtggagagtcaggctggcgagcggctcgctgtccactcggc
aatccaggatgcgccggaggccacctcaggctccacgaagaccatcatggtggggctcttggggaggtctgtggggagg
aaggaggtgtgtgatgacagaaaatgaccacaaaatttctccctccctgaccatgcacagtgactttgctgctccctcc
cattaaggagcgaacactatcttccactccttgaactctgtgactcagtttgaccaacttgagaaagaaagtgatgttgtg
caacttcagagccagacctcaaggagcctctgtagctctgctctactgttgtgaacacaaacatgtgtgggaagccagggc
agcctgctgggcacatgtggcccaactggcagccaggagcgtgcagtcactgcagagccactgggcaactgcagccacatg
aggtgtgaacacagaaagaaactcccagctgagccagccacagaaatgagcagataaacatgactttgttgaagcag
taagtgtgtgtgtcaatagcaatagctgactgaacagtgtaggcttggtgaaggctggagtgaggagaccaagact
gatggggcttatcgctgtgccagacacactcagcacgcttatttcttgggtttttctctcttcttcttcttcttctt
tt
ctccactcccaggctcaagtgagctctcctgctcagcctcctgaatagctgggattacaggtgtgtgcttcatgcctgc
taattcttgtatttttagtagagcggagtttgcgcataatgaccaggctagcttgaactcctgaactcctcaagtgatccg
ccacctcagcctctcaaggctgtgggattacaggcataaagccctgcaccaggcatttatttcttatacaacttgcagg
atgagccgaatgatgtattccccatttgcacaagggtcaaccaaaccagacagcgggagagccaggatccaacccaggactc
aatatccgtgaagccttcaactctggcacacagaagaggaggaagaaactgaggtggcggtgtgagatgtggggagactaa
ggcggaggagggggttcaactcacagagcacacggagcatgactggagcagaggcagcagcccgatggggagctcagc
ccaggcagtggtacatccagcttgagcacgagccacgtgggtgaaggcgagagtggggcaacaggctccgcgtgcagccgc
ggtcattccagaacctgcaaaggtggagttgcccacaggccaggggccaccaggaggcggcagctgaggttcaggggca
gcgccctcaggcagcgtccaggccaggctctgccaccacggcgtgcacctgcgggcggaggatagagagatgatggggatc
tgtaggccttgggggctggagcctctgggtgggacactctgaggggctctgcacattgtgggaaggtctcagctcgaact
ggtatagggtcttccaaggtggaaacctgcccccctccagtgggagctgtgcccccctccaccagattaggctctcacaaggc
agagcttctcagaccagacagcccgctctagtgaggagctccagccctctcgtggaactggacctgagagcatcagagc
ccctcctcctccctctcttccactcactcttctacctgcaacccgcccagtggtgctgattgagcccgcaagttttgggc
tgtgcaaacataggtgtcatcacctcagggcacacttgcaactgcagccgtgagggcgttccgggcaacctggacatggc
ctgtccctgatgccaagctgtggaccocctgtcgtgtggccagcaccttgccatcatgagatagggccagctcagcaggt
ggctcactgtccacagtgcaactgtatcacagcctggatctggccctggagtcggggaaggaggacaggacagcgtcctg
aggggacctctagtgggcaggcagggcacagatggggacacctgcttagggcacctgtactgctcatgtcctagctgctggg
gttggtcaggctgagggcctctggaccaaattggcacttctaagaagtgaacacctcccaggagtcctgtgggcagagctgct
ccaggcagggtctataggactgtctcttctctcagacactatgtcctcttccccaaactccccactcctggcccgactgc
agtcccttggccactccaccagagccccaagaccctccaccacctcctcctgggcagcctcggcgaaggacacctct
gctcacagaggacttgcaaggcagcaggagcgggagctgcgggtgccttgggactctgggctggcgaagataggcgct
gcataggccgcgctggccaccaggcgctggctgagggccagctggcgagctggacacctcctgcagcccaacgacgttgtgttaccaggt
atagagtggtgggtgctgggcagcagggtccgcacaggtcactgtgatgggggcacctcaggcaacggcagcctccggag
ccaggatcacccgcacacctgtgtccagaggccaggatcagccgggcccagcagcagcggaggtctccatccca
gtgggcttggcctaggccctgcttacctccagccgcagctccagggaagtggtggcctggccaggagggtgcggggca
gagcagctgtagaacctcatccctgggctgtggcctcgcagctccaggcgcagggtgtggggacagaggtcgtgct
cgaggggcccaagaggccagcggtggctgagggccagctggcgggcgggcggtgctccacagtcacagatccaggg
ccagctccggccactggctcctccaggaggtaggtcaggcgaggttgccggggcgctctgcagggtcagaggtctaca
cggagtcacaggcgagcagctctccaggaggccagtttgagggtggcattcaactcaggaggggcctggctccccacct
atggctcccaaacccagctcagggcactgctcctctgcccagacaggaactcaccgccagcccccctactcttaagtgt
ccttgccaggtgctccactgactaggttcccgagttccccctggacacactgcgggttccgggcatgcgtgatccc
tggggctctgctcattcctgtagtcacccactgcatccagccagactcacagaggacgtccaaggtgataggtctgga

FIG. 7

36/157

[illegible]

FIG. 7

37/157

ccacagagtgggttttgctgtattagatctctctcggagcagacagctccagagctccagccctgcccccagccaccat
tcctcctgctagctcaacctggccat tgaaggtggctgaggtggaggcttgcccacgggcat tgcctgacctcagcagaggt
caagccctctctccagcaaaaggttggtgaatctccacagcagcaagttgggggtcttggtgacctcatgctgggtggg
aacagccccagagtgctgcagcaacccccctgatggcagggaaagtggccacaacacgctctctgtggagcagctgcagc
ctggcgggggggtgctgtccacacgggcacaaagagggcctcgccctccagcccgccagcagggcctaaggttcag
ctctgggttgtaagtgttggtgtgcgaggggggtctgcagggagggaagaacatgggcactcatccacaggatgctccacgggc
cccaacagctcctgctaggctccccaatatacactgcacaaaggagcagctcccaaggctgccccpaagcagctgtgcagct
tgactgcacagggcgccagcagcagggtgggtggaaactcaagctaggctcatgtccccaagctaaagctgtgacctga
tgccctactgagctctccagaaagaggagcacaatttctatgcctcagcaaaaggtgctgtatgggctggctgccacttc
cataagctccactctgctgcgagcagggtggcccaactgaacaacctcgggggtctatgctgtctctccacctgccacc
aatgctctccacggagctccctcaagatctgcccctgcagcttattctgtgatctccaagcagctcctgtgatactag
cgaatgcataagaatggcctctctcctgaaggtctctggagacagcttgggccaaaagagagcaatcagccaaagccccatgt
gatactcactcagacacagcaaaaacttgacagaggcagctccctctcagacgttcagggaatccgggaatggctccgcaacatgtg
tcagggaagcctctgggaatcaggccctcttggtctttacacaggagcctccccctctggcgctcagagagggggctgtctgtcc
ctgtgctgaggggctttgctcttggtctctatgggagcccgacagcagctgtggggcattgtcaaaatagagagtggtg
tgtccagggtcacccatcagagcagaagaacatggatgtggtgcccattgctgagcttgaggctgggacagaggagccaca
gggctggcat tggagaggtaaagacatggggaggtgagagaagggcccccagctcatttccataacttcagcccccagctggaa
agaccactgtggccaggtctgtgacacagccctccaggaatagaaacttggggcaggagggagcagggctggggggcctta
gaggaggtggggctgcttggtgggtccggcgaggagggtccaaacaggtgggtggctctccacagagcagctgagaaacagct
ggcgaagaggggccactggcactgtggcctgcggggccggcgaggctgattgagccggcctcagctgtcggaggggccggg
gagcaggagaggtctgcgggggacctctgtaagcagggtccatcaggtaccaggagaaagcgggcatcaggtgtggggc
ttaggccgctctgcagctcagtgtaactgctgtctctccaccactcggctcggggctgatggaggagacggggcgct
cgggggagaggagaggaggtgggaaggggtccctctctcaaccacatgctgacctatagaagcctccagggcttct
ccttcccccactactgtatgtagtcttgggctttgtgggtcactcaggaggggtttgcccctttaatgcnaactcagatct
atttggtatagatggctgcagctggttggaactcagagccagagccagggtctatgtccaaactcggcttggtggg
catggtagctcatgctgtaattcttagcactttgggagactgaggcagaaggatcgtttgaggccaggaggtccagaca
cgttaggcaactatgtgagactccatccaaacaaatttttttttttgagactgagctcaactctgtctccaggctgg
agtgactgtgtgtgatctcggtctcactgacacactcgtccacgggtcacaagcatctctcgctcagctcccaagta
ctggggtatcagggtgtgcacacaccccagctaatgttatatactgtagtagagatagggtttcaacatgttaggcaggc
agatctcgaactcccgactctggtgatccaccgcctcagctcccaagtgctgggatcagagctgtgagccactgtg
gcaagcttttttttttttttttttttttttttttgagacggagctctgctcgtcagccagcctggagtgagtgggcga
ctcagctcgctcacaacactcgtccctccagggtcaagaaactatctcgtccagctccctcagtagtctgggacacag
tgtcgccacacacacccggctaatttttgtatttttggtggagatgggggtttaccacgcttgggcagggtctctcaagc
tctcgactctcaagtgtatctgctcgtcggtctcccaaaagctctggaattacaggtcagtgacccacatgcttggtccaca
tttttttttttaattagctgggtctggtgagctggggcagctcagctatcttggaaggctgagatgggagagtggtct
tgagccaggaggtttgaggctgcagtgagccatgaacataaccatgcactcgggcctgggcaacagagcaacacccctatc
tcaaaaaaagaaaaagaaacccgtggtctgataattagctaccatgcctcaggagagggaaggacagctgcataccg
aaagtgggaagcagctactttcttt
aggctggagtgagctggcctactctggccccactacaactcacaactcctgggtcacaagctattctcgttgctcagctc
ccaggtagctgggactcaggaaactccaccactgcccaactaattttgtatttttagtagagatgggatttcaacatg
ttggccaggatggctctcgatctcttgacctctgtgatcccccactcggcctcccaaagtgctgggattacaggagtgag
cacccggccacccagagaagccactctttctatttggtctccacatctgaactctagcatagagctgctcccgag
ttctataattaaaaaacataaatgtctgtgggactcccaagtttaggtcagggttaggtgcaaaaggaatgatgtg
acctgtgtgtgtttcccttttctccttgactccaggaaagctctgcgctgggtcagctgcacagccctgtcttttattgt
gcgtagccagcttagctcagctcgtcggttgagctcagctagactctagaaggaaacagctggagcaggtcctctcagc
ctctccactctccctggctgcagctgcccatcttaccattggcatggaagtcagggttgagggttgcatttccaagg
agttgggtgctgagcacttgtaactcccatctgcagtttgctcaggtctcgatctccaggcgcaggagttgggaca
gaggttaactcaggaagctgggtctgtactgtccocaggagtgaggccaggatgatagcccactgtgacagacag
tgtggccaggggctcactgaccacagagcagtgaggaatgccacaagctcccgctgggtctccaggaaagctgtcaggga
ctggagtggagggcggtctcgtggagacagaggtgggctgtcaacctcaagacaagggaactcctggaacctga
tacaacctggaactctgcacgctttgtccacactgcctcgagaaaggggggtgatccaaagtgctgagtgctgc
tgtaatacacttttgccttggtgggtgggtcagctcagctctgtgtccagctcaggtgtcaaggccacccctgcagggtg
ggataaaaagccccactcaagagtggagaaactaatagagagactgtggaggagacagactcagccagcaggtgaccaaa

FIG. 7

38/157

tggagcatcatcagcaatgggacaagctgacacagagtgccctcctgacagggcagagcagcatgtccgcgccctccacc
cagaggggcatggtctcaggccattcaagtacggcaaatgtggagtgagggaacctcctcacacagcaggccctgcactctg
caagtttcaacgtcaagaatgacaaggaaagactgaggaaacagtctcagactaacggagaatgaagagacgcacacgccc
aatgcaatattgtgaactgtgattgtattctctggaccagaaaaaaatggctacaaaagacagatttaggtccactggctaaa
atgtgaatatagattagcttagataaacgtcttctatcagccagggtgtagtggtcctacacctgtaatccagctacttt
gggagcccgaggtaggtggatcagctgaggtcaggaggttcaagaccatcctggccaacatggtgaaacctgtctctaca
aaaacacaaaaattagccaggcaggatgggtgggtgctgttaatcccagctactcaggaggctgaggcaggaggagaatcg
cttgagcccagaggcggacattgcagtgcagcaagatagcaccattgcactccagccctgggcaacacagtgagactcca
tctcaaaaaaaagctcttctatcaatgctaattttcctgatcatgatcattgcattgtggatatgtaagagaatg
atcttaggaatttagcggtaaaaggccctcatgtctgcaacttcaggaaatataaatagttaggtagatacataaataaca
tatgcataatgtatagagtgcccatatattgtataagtgcaacatgtccacatagagtggtatgtgcatacacacaagtga
cctgtatataatgcaagtctatatccatcacatttatattgtatgtatgtgcgtgtgtgtgtgtgtgtgtagacagcgagaaaga
ttaagcaaatgtgccaaaaattgtaataaagggaactctagggttaaagggtataataatcattatttctctctgcaac
ttttcataagtttaaaccttccaaataataaaggagggaactaggagggtcaaaagcatgcccagggtgagcagagctgg
gacatgaggtccaggccacctccaggtaagtggccttgaagttcccattgcccaggaccccaacctcctctggggctcca
ccatcctggctcctctgactgtgtgcagtgccaccccaacctggcctcttaccggttgactacacagctgacagggcccg
agcgctcgctgccatggaagcttctgcacctcacagaagtagaagccagtagtcagccctagtggccaagtgcagccggagg
gtatgggagtgggcatcctccagcaggacagctgttctgtaccagctgtagcggagctcactgggtgctcgttgggtgt
gttgacagactagtgctcactgtctgggtctctccaggatgggacctgctgggctcaacctggacctgcagccactgcagggag
cataggagtgctgggggggtcccagccaactccagccccagccccatcagctcccgggtctcagccaatcagccctcaa
tctctgcacatcctacacaccccgctcgtgctacagggggacctcctggagtgctcgcactccttctgtcctgctcacc
aggatccccctgacacccccacctctgtgggggtattgtgcacctgtctcatggttaaggcagggtgggacctcacacctg
catccagggaagcactgggttatgccagtgctgtggcctgcctcctgctcctcctcctcctcctcctcctcctcctcctcct
ccctattctctggccaatagccaaccttcccagaagctcactgaagatgtggagggtgatgggggggtgagaccaaagagc
ccaagcgttctcagcttgccaggtgtagacgcagcatcgtctccaggctgctggggcagggtgcagcacacacagctcttg
tttgagggggtccccctccttgagccaactaatggaaactgactgcagggtagctgctgttcaacctggcagggtgagttg
gaccagctcaacctgggaaggtattcctccccgaggggctgaggaggatcttcacacccccgggggcatctgcaagtca
gtaggggggtattgggtaagtgcttggggaggggcagaggatggcacaactcttcttgcocccctttaaagctcagctcct
aagggaagtatgccagataaaaacagcagctccccctcaatcctcaccaggggcatgctctgtctgcccatgtcgtcctct
tcccgcocctcgcctgatgcattcctatgcccatgagagctgatcatggaagcttgccagcgtccagcctacccc
accagctgtagttttctgcttcccaatgctgtccattgctcctgctctatttgggatgagctccacacacaggggtcatgg
gtaccccactgctagcttttagtggcctgtggaagctatttggaagggaaccaactacctagtgaggagggggccaaggcag
catcaaacctagctctgaaatagttaccagtttgtgaagcaaggaggccaacaacacaaaggaactgcatttcttaaatct
tgttccaaagcctcctctgtgtattcaagtgttttatctctatttttttagaaaggaggtatctggctcagtcagccaagc
tggagtgacagggccaatcatagctcactgcagcctgaaaactcctgggctcaagtgatcctcctgctcagcctccoga
ggaaactgggactacaggtgcagcgcacacacatccagctaaattttgttttttaattttttgtagagacagggtctcact
atggtgcccagggtggtctcgaactccttgctcagcaaatctcctgctcctggcctcccaaggcactgggggttacaggt
gtgagccactgtatccggcctcaagtgtttaatttggccaggcaactcttaaatccttgacctgggtcatctccttta
ttgtttttgttgttgttgttgtttttgttttgagacacagctctcactctgtcaccaggcttgagtgagtgagtggaaca
cacaaactcaatacaactccaccccggggttcaagcaatcctagtgctcagcctccttagtaactgggatcagagcat
atgccaccacacccgggtcaattttgtatttttagtagagcagggtttcaacatgttgcccaagctgggtcttgaatccc
tggcctcaagtgtaccccgctcagcctcccaagtgatgggattacaggcataaggcaccgcggccggccatctccc
tttaatttttatcataaactctgtagagcaggaacctctattgttatctgtcagctgcagaaaacagagcaggga
ggaaagggaataaatcccaactccccagcatcctccagcagcgggttaaatcccggggcaactgtgtgccaacccctgaagctg
catggagctcctgttccccaccccttgtgtctcctcttctctgcccagcctgacctcaaggtaacggccttgcagtc
agtaggaggttgagggaacatttaaacgaagtccttggggctcagcagtgctctcctcaacttgccctacacctggaggcca
ggcagcagcctgggtttgcatcaggagcaagggttcogtttctgtgggctggagagggggctgggtttctgtcaggagcaaca
gatgcgctcagccacaagggtgtgtccccagcaactcacactcacttgaggtgaactcctcgtctgagccctgtgtatt
ggccacggagagctggcagcgcaggatccggcogtggtcctgccaagacatggccatgtaggggtctccaggtggcoga
cgccgggtgggtccaaactcttgggtcttggaaggttgaggtcctggccttgccactcagctcagctctgacctgc
tctctcaggcatacgttagggagtgagcagttgaagtcacactcctgctcctcgagaagctccacgggggagcaactggt
gggacacctgggtcctctgaggacagagacagcagtgctcaggaccogctttgcaacccctgagcctcctgcctggg
aaacccacagctgaggagagagccctggggagggggctttagggaaggaaagacgctttctactccagcccaactggg

FIG. 7

39/157

tgaatacaaggagaaaccaggcctgggcctacgcccgggctggaaacagaggctgagactggctgggggttagattcaggac
aagggtcggggctgagagccaaagggttcagaagcagcttgggaatccctccggggggcagccaggccacccacttat
cacctgttactgtgaccaaggttgcccttccacatctgaccagcgggtgacacctcactgactctcgaagcgggaaggtttaggaa
ccagagctcctcgggctgacaggtccttcagcagcaggttgacacacctgtgctcgggggttccccatgaactcgggtcggcc
gcggaagcggggcctccaccagcttgggggtccgcccagatggctcaccacctgcccgtgcccggagtagctgtagtaccaga
tggcgcgtgatgccgtcggggcactccacgtcggcagggaagctgaagatgcaggggataagcaggcaagaccccttca
ccctgcacgtcctggggactggagacgcccactgagggcctggcctgggggaagaaacggcaggggggaacagaggggaggggtg
atacaggcctcagggtgccaacagagccctgacacctgcccagagaagggtgcccagctggggtccccaatctctgcccctg
ccccagaaaaatgcacacttagagcagcccttctcagtgccccagggttcagtcacctgcccagggtcctcagaggtttg
gtagggtggctcgaagacagatggcagcttctctgcccagctcctggccactgcccagattggggcctcccttgacctcagc
aaccaaacatgtgacccaggcctagcttaactctgccaagggtggacttatgaaggctggaacccctggaagacaggcacta
ggcccgaagcttacctgctggggaagatgaggccaggaggagaagcttgggcaagaagcccatagcaggttctctgtgtg
tgctcctgttgcttaagaggggtggtgcccactgcccgtggcctcagcaggggcctccagggaacacctctgggcacttt
agcccagcacctgctagaagtcgagcctgtgtccccacctcctcctgctggccaaaccaataagaggggcagggtcctta
aagacctctgagtcagacaccagcagagagccaggaggccacgttcccagctcaggctgtgtgcccaggaatgcccctcact
ggtggcctgctcagaaaaagcctgtgtgtcctctggccctatccccagttctgtcttcccagccctcaaggataccacc
ctaaaggcagatgaacagctgtttctccctcttccccacttccctgcccctccccaccacccaaacacagggaactgga
gcccagagatgcccagttactcactccaagacaccagctagaatgatgggttctctcctgaggtctgtcctctaccaact
gcttactaactatagaccataatggggcttactgaaacttgccgaagtgtgcttcttaacagagctccctgctcaaa
aaccttctgtgctccccatggccggtgagatgtgaaaagtcctcatctcctgcccagctctgtccccatccctgct
ctccgcagacaccactctggggcagcttctgctcagggtcctcctagctgcccagctctatctccaccacagataat
cttggctgctgaaaccttatcaaccttcaaggagcagcatgaatttggtctccagctggaactctcctcttgagggtc
ctgtagctacccagagctatctctatttctcctgctgtgttttacagctctgtgagagccacttctcaggacctaga
actgagatattgtgcccactagcagctggcaggacctaccaggcatcagcaaaccccttagtgactaagagaggggtgagg
tctttagggttccagagctcagggttcagagttggagttggggagggtggcagggaaggtctagggttgaagggtgagctgag
aggctgtggaaacacatccccacaaatgagctcaatgtgtgcccagtgtaacatctcaaaagggtgagccctcaac
catggagtgcccctggtcagggtggtctcaagggttaccatagacctgagcccaaaagggaagagatgccagaaacacata
taagtgaacctataaaactcttagaagaaaacaggtgaaaaatcttcatgacctggattaggcgaatgcttcttaagtat
gaaattaaaagcagaaaaaataataggttaagttggacttcatcaaaatttaaaacttggccagacacagtggtcctat
gcctgtgaacccaacacttggggaggctgaggcaggaggatcactttagcccaggagtccaagacagagctggggcaatac
tgcaagactctgtctctacgaaaaataaaaaaacaggcctgtggtcccagctactctggagggtcagggtgggaagatcgc
ttgagccagggggggggtcgagggtcgcagtgagccatgatgacacactgacactccagcctgggtgacagagcaagacc
ctgtctgaaaagaacaaacacagctgggtgcccgtgagctcagcctgtaacccagcacttggggagggtcaggcctgca
gatcacaaggtcaagagattgagacacctctggccaacatggtgaaacccctctctactaaaaatacaaaattagctgg
gtgtgtgtgtgtgacctctagctcctagctactcgggaggctgaggcaggagaatcactgaacccaggaggcggagggtc
acggtgagccaggatcacgccaactgactccagcctggtgacagagtgagactctctgtctcaaaaaaataaaaaa
aaaaggtatgaatcaaaggacactatccagagagtgaaataaaaggaggacaaaccaagaatgggagaaaaatatttga
aatcatctatctgataaaggactaatatccagaatatataaagaattctcaaatgaacaaacaaaccccaaaaaaac
catgaaatccaaactccaaaaatgggcaaaacactgaaataaacattcttcaaaagaagatatataactgggtgataaagg
catgaaaagatgctcaactcactagggcattaggaaatgcaaatcaaaacaaacacacagataccacttcaatcc
gtttagattggctattcacaacaaacaaaggcaacacagaaaaacaaataaatttggtgaggatgcgaagttgaaatctcttg
tgtattgtgtgtgggaatataaaatgggttcgctgactgtggaaaaacaaatgggtcattcctcaaaagtcaactaggat
taccatattctccagcaattccactcctaggatatacccaaaagtactgaaacagggaactcaacagagtcacacaa
gttcaaggcagcactattccactaaaagggtgaaacagggtcaagtgctcactcagctgaatggatggaataaacaactgt
ggatgtacatacaatggaatatcaatggcctgaatgaagggaatgaattctcaacatctgtaacctgaaacctgaaacatctg
tcagtgaacccaggcagacacaaaagggaataatttgagggttccaattacatgaaatctctagactacgttatattcag
agactgaaagtagaattaggtatagggtaaacagggggtgcagggaaggggagctaatgtttaatgattgctgagctctc
agataatgaaaaagttctggaaaatagtggtgatgtttacacaaatattgcaaatctgaactgggggtgtaacctt
aaagacagctaaaacagtaaaattttacattatgtatatttcaaccaatcacacccactgtgccaacttggcaatcctcc
ctggtcctaatactgctgacttggcccaagctgaaacaggaggtggaagtgggggaagggaactcagacctccagctccca
gcccactccgctcagggggtgggctcagcagattccaatgactaccagggtcacacctgggaagggggtgagccgaggc
ccaggggcagtcaggctgaccaggtgggacttagcctgctgcagaaggcagaaggtgcccagcagggggacagtcag
ggcgggattgggacagggaaggaacacgctccccagggaaccagcctctcgcaggctctggagtggaactgatctggcc

FIG. 7

40/157

at t t a t g g a g g c c c a a g g g c t c a t c t c a g t t c t c t a g g a a g c c c t a g g c c t c c t c c t c t t c t g g g a a g a t g a c c c c c a
g c c t c c a c a c c a g g t t c t g g c c a c t g g a g a a t g a t a t a g c t g g g g c c c t g g g a c c t g g a c a c c t c a c c g t g a a g a a a g
c a g c c t g c t g g g c a c a c t g g g g g t c a g a t g t g t c c c t g g c c a c a g g g a t g t c a g g g t c a g c t c t g c t a t g g c a g g g c a
g c t a t c t t g t c c c a g c t c c c c t g t t c c t c c c a t t t g g g g t c c t g a a a a g g g c a a c t g t g a a c c t g a t g g a a g a t g g t g g
g g t t c t g g a c a c a g c g a c c c t g g a a c a g g g c g c g g g g a g g a c c c t t t c a g g a c c a c t c c c a t c a c a a t g t a g a g g c
c a c c t a t g c t a g c c g c c c t a a c c c c a a g g g g t c a g c c c a c c g g a a t c c a g c c t a t t g g c t c a g c c t g a c c a c c a
a a g g c c a g c t t c a g c c a g a t a a c t g t t c t g g a a c a g a a a g a g c a g g g a c c g c t c a g a a a g g a g a t c t c t g t c c c t g t t
t g a a a g c c t g g a g t t g a g g g g a c a g t g c c c c g c c c c c g c c c c c g c a a c t g g g t t g c a g c t g t g g c c t a g t g a g c a g
c a g c g c c c c c t g g t g g t c g a g g g g g a a t t g c g g g t c c c g g g a a a g g g g c g g t g t g c c a g a c a a c a g g g a g c a g g c a g c t c
t g c a g c c c t g a a c c a t c c c t c c c t g g g t g a c t c t t t t g g a a a t c a t t g t t c c c c a g a c a g g a g g t t c c t g a g g t t c a t a
c t t g g g t c t c c a a g t c t t g g g t g c t c t g a a g a c a g g a t t t t a a a t c c c c a c t c c t a c t a t t g g t a t g t g t t g c a t c a g g g
t g g c t t g a g c t g g c c t g g t a c a c a g t g g g c a c t c a c g t t t g c t g c c t g t t g a g a c c a a g t g c c t c a g g a g g t c t g g a g
g g t t g g c t g g g g c c c a a g t c c c t g a c c t c t g a t t c c a g a g g c c a a g t t a g c t g g g g a a g a a a g g g c a g a g g c a g t t c
c c t a t g g a c a g c t a g g c c g g g t g a g g a t t c a g t t t c t g t t t c c t g a c a c c a a g g c t t c c c c a a c t t c c c a t t g g g
c t a g a a g g a a g a c a c a g g t g a c a t g g c c a g c t g g a g g t t a c t g g c c c a c a g a t a g g g a g t c a g g g t a c g g a t g g g g c
a a t t c t g g a g c a g a t a t g g t a a a a a t a g t g g a a a t c c c c a a t c a c a g g c c a a a g t t a a t t c t c a g a g c c a t a g a a t
c c a t a a c t a a t g c t a t t g g c t t t g a c a t g g g c a g t a g a a a t t t c a c a c t t a c t c c t a a t c t g g c t c t a a a t g c t c t
g g c t g g a g t g c t c a c t c t c c a a a c t g t c t g g a c a g c a c c a g a a g c c c t c c t g t g g a c g g a c a a g t g g t g a t g g a t g a a
g t g g a a t t g t g c t g g g g t t a g a t a a g a g a g a t t a t c g t t g g g c t g g g c t g g g g t g a g g t c a g g g t g a t g a t c a g g t
g g a a t c t g g a a g c a a a t a a g g c t g g g a t t c a g t t c g t t c c t g a c a c c a a g g g t t a g c t g a g g g t g g g g t g g g g t g g g g t g
g a g t t g g g t g c a g g t t a a g a t a c g g t g g a g g c t g g g t t a g a t g a g g a t g a t a g a t c g g a g t t g t g g t t a g g g t
t g g g a t g a t g g a t g g t t g g g a t t a g a t g a g a t a g t a a a t g g t t a g g g t t g g g g t c a a g t g a g g t g a g g t g a g a t a a
g g t t g g a a t a g g g g c t g g g g c t g g g g c t g a g t a g g g t c a g a g c a g g g t g a t g g t c g g a g c t g g g a t t g g g a t c a g g t
t g g a a g g a t t t g g g g g t g g t g g a a g g g t t c a g t t a g t c c a t c t t g a t a g t g t c t t g g a c t t g g g t t g g t t g g g g t c
c a c c a c t g c a c c c a g a t g g a g c c c c c c c g a c c c c t g c c c c a t c c c g c t c a g c c a g t t c a g c c c a g c c c c g c c t c c t
a a t g t c c a c t c a c c c t c t g g g c c a a g g c a c a g g g a c a g g g g t a c t g c t c c a c a g a a g g a a g t g g c t g c g g c g g t g c t
g g a c c t g c g g a g a g g a a c a g g a a g g a c g g c c a a g a g c t c c t g g t g c a g c t g g c t c c c a g g g c t c t g c c g g c t c a a g a
g a g a a g g a t c c c g t a t c a g g g g c t g c t c c t c t t c c c a a g c c t c a g c t c t a c t g t c c a a c c a a g a g g c t g g t c a g g g a
g g c a g c t g c a g g c c t t g c g a a t g c c a a a c g g g a a a g a c c t c c a t a g g g g a a g g c c t c g g c a a g g c c a g g g a c t a g g
g a c t c a g c a a g c a a g t g g g a c c g c t g c a a c g c t g g a g c c t c c c a g g c c a a a g t g a g a a t g g a t g g g g g a c t c c c a
t t c a c c a g c c a a t c a c a c c c c a c t c t c t c t g a c c c c t t a g g a g g c t g g g g g a g g t g g a a g g g g g c t c c t g c a c a c
a g c t c t c c t c a c c t g a c a c c t g a g a g g g a g g c g c g c c c a g c c t g g g g t g g g g a t g c a c t a c a c g a t g c c c a g a c a a g g c a
g t t a t t c c a a g c a t t c a g g a g c c t c c c t g t a c c a c t g a g t a c t c t t a a g a c c c c c a a g a g g c a g t c c c g t t t g t g
g g g g a c a t a a g g c c c a a t g t c c a g g a c a c t c c t c a g g t t t t c t c t t c c c c t c t c a c t g t c c t g t a c g t t t t t g g g t t t
t t g g c t c t t g t g t g g t g g t t t t t c t t t t g t t t t g t t t g t t t g t t t t g a g a t g g a g t c a c t c t g t c g t c a g g c t g g
a g t g c g g t g g c a c a a t c t a g c t c a c g g a c c c t c c a c c t c c c g a g t t c a a g t g a t c a t c t t g c c t c a g c c t c t c a g a t a
g c t g g g a t t a c a g g c a t g a c a c c a c a c a t c c g g c t a a t t t t t g t a t t t t g g t a g a c a g g g t t t c a c a g t t g t g c c a
g g t t g g t c t t g a a c t c t g a c c c c a a g t a a t c t g c c t g c c t c g g c c t c c c a a a g t t c t g g g a t t a c a g a g g t g a g c c g c t
g c a c c c g g c c a c t a t c c t a c c t t c a c c c c a c c t t g g g a c a g g a a a g g a a g c c c c c a c a c a g c t a g t t a c t g t t a c
a t t a c t g a g c c c a t c t a t t g a g c c c c t c c t t t g t g c a g g c a c t t g a c a t g a a g c a g a t g t t a a t t g c t c a a c t c a a g c a a
t c c c a g a t a c a a g c a a a a g a c t t t t t t g g t t t t g t t t g t t t g t t t t g t t t t t g a g a c a g t g t c t t g c t c t g
t c g c c a g g c t g g g g t g c c t g g t g c a a t c t g g c t c a c t t g g c t c a c t g c a g c t t o g a a a t t c t a g g c t t a a g c a a c t
c t c c t g c t c a g c c t c c t a a g a a g c c g g g a t t a c a g g c c t g c c a c t a t g c a c g g c t a a t t t t t a a g t a t t t t g t a g
a g a c g a g t c t c c c a t g t t c c a g g c t g g t c t c a a c a c c t g g g c t c a a g t g a c t c t c c c a c c t c a g c c t c c c a a a g t g
c t g g a c t a c a g g c g t a a g c c a c c a c c t c c c a c t t t t t t t t t a a t t t c a c a a c t c a a a c t c a t a c a t c c a c c c t t g
c t g c c t a c c t g t a c t g g t g g g a g a g c t g g c c a c a g c a a a t t a g g c c c c t g t a c c c t g a g g g c a a g g c c a g t g t g a g g t g
t a a a a g g a t g t a a a c c c t a a a t c a g g g t g g a t c a g a a t c g c a g g c c a t g g t g c c c c a a a g c a g a t g t c t g g t g a c a t t
c c a c c c t g a a a t g c t a g g t c a c a g a t a t t a g g t c t c t a t c a g c t g t t c c t c t t a a g c t c t g t g c c a c t c c t a
t g c t t g g g c a t t t c c t c t t g c g g c a a a a c a a a a g g g t t c a t c c c a t t a c t c c t c c t c a a c a g c t g t c c g g a g a c
a c c a g c t c t a g g c c t g t g g g g t t g t g a c a c a t g g g a t c c a a t c c t t a g t c c a c t g g a c t c a t a c a t c c a c c c t t g
g c t t c a t g g t g g g a a a g c a t c c c t g a c t t t g a c c t t a g t c a t a t g a c t t g c t c t g c c a a t t g g a t g t g c a c g g a c t g a
c a c a a a c a a g t c t g a a a g t a c t a a g c a g t t t g c c t t c c c c t g a a t t c t g t c a t t g c c a t g c a g t c a a g g g c a g g c t
c c a a c t a a c c t g t g g t c c a a a g a g g a t g a c g a a c a c a t g c a g a t g a c c t g a a t c a g a c c c a t g g a t t g a a c a a a a c t c

FIG. 7

41/157

agctgagccagcctacgtccaccagaccagtcaacctgtggatgcgtgaatttattgtcggatgctgctgagaattttgt
tggtcatttagcaagatgatacaaggcctaagtcccagaacaacaccccagaacttgcttaacctttccttagcatgag
gagagcaaaagacttgtctaccttgattagtcaggaggcactgcttctgtcatttcccttgagtatacagcaaaactaggt
aaataaataaaaaataactaggtaggctgggacaggttggtcaccgctgtaattctcagcactttaggaggccgaggtggga
gatcgcttgaggagccaggagtcaagaccagcctgggtaacctggcgaaacctgtctctacgaaaaatacaaaaaattagc
tgggcctgggtggcaggcgctgtagtctcagctactcaggaggctgaggcagagaaatcgattgaaccgggaggtggag
attgcagtgagccagatcacactactgcactccagcctggatgacagagcgagactctgcctcaaaatatttaaaaa
tgtaatttctcagtaggtcacactgttacacattctctaataacaattattcttttttttttttttttttttttttgg
agatagggtctcactgtcaccaggctggacaggctggagtgcaatggcacaatctcggtcactgcaactttgacctcc
taggctcaagtgatctcctgctcagcctcccaagttgctgggactacaggtgagtaccacagacgcagccaattttt
gtattttttgtagagatggggcttcacctatgttgccagggttggtctcacactcctgagagtcccaagtcaagtgtacc
accaccttggtcctcccaagtgctgggatcacaggtatgagccacatcccaagtgaaattattctcgttccagatagga
aaaccaagctcatagggggttagagaatttgccaaagacaaaaacttttgggttgaaaaaaaataagttttgttacaagta
tagaaaaacacaaaataacgggtgttttaataaaaatagaagtggttccacctctccctcaagtaagtggtggcatccaaga
tgatatgacaactccacaatcatgaaactagatccctttttttttttgggtcagtcattgtcaatgggctgctccagt
cattgtccaaagtgggtgcttgcgctccagccaactgtatctgcatcogggaagcaggatggaggaaaggacaaaagtagg
ccatgacctctactttaagataactccttaaattgctgatgtgggtgggtcactcctgcaataccagccactcaggaggct
gaagcaggaggttccactgaaacccaggagcttgaggctgcagtgaaactaatttgctcactgcatccagcctgagtg
cagagtggatctctgtctctaaacaacacaaaaaggtaattctttaaattgtacactttattgtcgttatattccac
tagaagctagaagttagtacatgggtcttaagtccttattctaacagggtgagtgcgccagtgaaatcctggggggccagt
gctaagagctagaggagatgaagcctcctcctgtggagcaaccaggacotttaattcagtcattcattccttggaac
taccagatccagcctcaggccctcggccacctccttgccagtatctctgcaagagcctcctgcctcccccagagggtg
cagagccctctgcttctcagcttcagacccacttgctgtgtctcaggtggggacaactcagctctatagaggccaaagg
gagatctttgagaccatctgtagtacgtctggggcagatgggagtcagtgtagctcactggctaaaagcctaggtcctg
catlagagctgggttcaaatccaggacatccactcctgctctacgatccaggagtcactcaatgactcagaagctaga
gtcctctggaaagcttgtatgaagagttccaggggcctggacaatcaaaagagctctgagaattgtgggtctgtgctgtgc
ttctcttattgggtatctgaaggcctgactctcctctcctccacccagaggtgacctctactctcatagccacacccctg
gacctcactcaactcagtcacctcaccagctctgtctatctgctcctctgggtctcctctcctgttcatgcatcttgtact
caccactctctgtccctgagtggttcaaggctctgggggttttcagagacactgaagggaacctcctcctcaacacaccca
caaggtctagggtggaatgaccactaagggaacgggctctgaagccagagactccagggaaggtcaacaaagcccaaggt
gcccgttacaagaaagttaaaagaccatgtacatgtcctccgtttttattcctcctcctcagggtctgggcatcacagtga
acacatgcagtcccccaaaggacacgctctgtacagagtcagatggagttaaagaacttagccgggtgggcccgtggct
cagcctctagctccagctctcgggaggttgaggcaggagaatggcgtgaaccggggaggcgagcttgcaagtggcgcg
agatcgcgccgctgcactccagcctggggccacagagacagactctgtctcaaaaaaagaaatatttagcttgagccg
gcacggaaagcttatgcctgtaatcccagcacttggggaggctgaggtgggaggatcaacctgaggtcagaagtttgagcc
agcctggccaacatggtgaaacccatctctactaaaaatacaaaaaatagccaggcatctgtaattcccagctcactcagg
aggtcaggcaggaataacttgaacccaggaggtggcgggtcacagtgagccaagattgtgccagtgcaactcagcctg
ggtagacagacaagactccatctaaacacacacagcgcacacacatatttaaacccacacacacacacacatctagtcca
gatgggtggactgagaacttgtctctgccattcctggcccatccaataccactgagagcacagtaagcaaaagggaaggga
gacagaagggtcgggaacagctggctgggggagtggaaggtatccactgcacaggtatttgatttaattcagaagatag
aaagaggaggagatcacgttcaggaaacagtgccgggtaaaggaaaccagagccaaagcagctggagagaagctgccccag
agggccgagcagaagtgactctctacaggagctcaatcacccccaaagggttggtagctggcggacagctacctctccacc
cccacatgtaacctgcaggggcagaggaaataccctaggtgagttccagttatcaatcaacctgacctttgttcataaaga
tgaattgggttaccaggttataccagacaggttgaggaaagacacacaaagagaaagatcagaaaaaaacaggccaggcg
cattggctcagcgtgttaaccccagcactctgggaggccaaagtggggtggatcagctcaggtcaggagttcaagaacagc
cttgcccaactgggtgaaacctgtctctactaaaaatacaaaaaatagctgggtgtgggtgggggttctcctgctcagcct
cccgagtagctgggaggtgaagttaggggttcaactgaaacccaggagctcagggctgcaattgaactatgattatgtcac
tacattccagcctgagtgacagagtgagatcttgtctcttaacaacaacaaaaaggtaattctttaaattgcacactac
attgtcgttatttccactggctagaggttagtcacatgggtctcatgtcttattcttaacaggtgagtgcccgagta
atcctggggcccaagctgtaagaagttagggggagatgaagccctgcctcctgtgaaacccctcaggttagaggtgcag
agagcggagatcatgcactgcactcagcctgggtgacagagtgagactccactcaaaaaataaaaaataaataa
aataaataaataaataaataaataaataaataaataaataaataaataaataaataaataaataaataaataaataa
ttagacaactttttttttttttttttttttttggagatggagttattttctgtcaccagggtggagtgaggtgggtcaa

FIG. 7

42/157

tctcagctcactgcaacctctgcctectgggttcaagcgattctcctgcttcagactcccgagtagcttggtatttataggc
 acatccaccacacccggctaattttggattttttagtagagaccgggttacaaatgggtttaccatggtggccagctct
 agtatcaaacctcctgacctcaggtgatccaccccgctcagcctcccaagtgcaggat tacaggcatgagccaccacac
 ccatccagtggttctgacaattttatacttattaaactcatttagcaacctctgagctagatgctattgttatctccactt
 acaggaaactgagtcacagaatggtacagtaaaataaccttgaccgaggtttacatggctggtaagtgccagagaaatga
 aacttgaacccaggcaacctgtttcctgagtggaactctgaactactcagccatactgcctatttatttatttattt
 tacttacttacttatttattgagatggggtcagccgggcatgggtgcttacgcctgtaatcccaccactttgggagggtg
 aggggaacagatcacttgaggccagcagttcgagaccagcctggacaacaggagacagggtcttgctctgtcaccagg
 ctggagtgcaatgggtcaatcatgggttcactgcagcttcgacctccaggctcaaggcatctcccacctcagcctccca
 agtagctgggaacctaggcaccactatcatgtccagctaatcgaaaaaaatattatatagagatgggggtctcacta
 tattgcttagactgggtctctaatcctgggcttaagcaatccaccaccttgccctcccaagtagtactagaattacagggtg
 tgagccacgtgctgacaaatacttctctttaaataaactaaacaaaaagtgaaacaaacaaatcccggggggaacagat
 aattcaaggaaatacaagaaaacttatgaaaagaagatcttagtgcaaatatgtttcataaaaaaaagagcagctatg
 aaaaagaagcaaaaaatattcttggaaatttataaataagaatttgctgaaataaaaaagaattcaatgaaaagggtgtga
 acatctgtgggttttgggttccaaaaattagtctcactctcttctgggttaaaaaatccctgatttctcgtgtatttctcaag
 cctgtgggttgggtggaattgacacctcttccatttactgcacagagagaagcatgtgtgctgctcagctcatttagagg
 agccagcaggttaagacaaacactgagggcagatgaagatatatactgacctgggtcctctgtggacatcttaggtccc
 tggatcaagccttaacctgaagctaaaggagtagtgcttctcactgttaataggcactgatcaaatccatagaaggact
 ccaattatccttgctcagttgtgtgcagcctcggtgagctacctcaagggaactgcccagaccttggtgacaag
 ccccccactgaaaatggggaaaaaacccaagtcacatcattgaactgaaagagctgcagttccccgaagacagggagag
 aagggaatgctggataaacacaaacaaactctcactcactgcacctgcaaaagaaatccaggcaacaggaggtgggacacagc
 aagtagtcttctgggtgacataacaaagacaacatctgttcttctcagaaagaagtaagaaggcaattggagaac
 cagaaacagacaaaatctatccaggcaaatagaagaacacacatctgaagcatgattttgaaacaaataggctgggtgt
 aagaagaagagaatccatttgaccttgaccttgaaacagcaaatattctccatcaaggcactttagggttagaggaaatgta
 tactatttcttctctcaaggagctatctgggttacctgggttctgcagtaaatgacgtttatacaaatcataaacctgggt
 ggttttcaatttctagaattcaacctgagatgaagatctggaagatttcatataatccagcaacagcatgcaaatgttac
 aaactttgacctgaaaatgtaaaaaagacccaggcgagtggtcactgcctgttaattcctagcacttctggaggccgag
 gcggttggtatcacaaggtcaggagtttgagaccagcttgcccaacatagtgaaaccccatctcactaaaaatacaaaa
 tttagctaggcatggtggcagcagcactgtagtcccagctacatgggagggtgaggcaggagaatcacttgaacccaggagg
 cggagtttgggtgagcaagaacacgcgcggcactacagcctgggtcaacagagcaagactccatctcaaaaaaaagaa
 aagaaaaagaaaaagtagacagcagaaattgaggggagatggcagggtgacctggggctggggaggcacaattttct
 ttacatagtgatggggtcaagaaattcttctcaaaattgattgtagggaacgaatatttaaatatagattgggttagaa
 ctataagtagcaccaacacaggaactccagtaatacacacagcaaaacaaactgagaaacacagacatggggcgggagatggg
 aatgggtgatttccattccctgctttaatgatgagggtgtcagtagatgctctcttgagttactaaattgagaaacaaaga
 taaagtataattggttagaggagtgatgtcaaaactcagaaaaaggtggcgcggtgggtcactacacctgtaattcctggcac
 ttggggaggcgaggtgggcagatcacttaagccaggagtttgagaccagctaggcaacatgatgaagccctatctct
 actaaaaaaacaaaaaatcaaaaaattagcaaaagcatgggtgggcacacctgtagtcccagttattcaggatctcacta
 ctgtactcagcctgggtgagagtgagacctctctcaaaaaaaacaaaaaaacaaagaagaagaagaagaaga
 actccagaatataaaaaatgaaaaggggcggaagagtggaactgcggggttagaaaaatttgcattttggtttcaatctct
 ctgtatatgtgaacttttttttaaccttactgactacgtcatatccacagcccccacacaggttaccaggcatcaact
 gctggttttatatggagggttgagcaatcattcctgctccttccagttcgtcttctgtactgcagaggctagaaaaact
 aaatttatataccccagatctccctccagctaccttaatgctaccacctcattcagccatcattgattctccactcacc
 aaactgggtcaaatattgtctcttctaaaaatttagaattgggactaaagagacactcagctgacaccttaaacacactgg
 gtttgaaactgtgactgctgcttatacacagatttttaaaaaataacacatatgaaaaaaatttggagagatgtgaca
 atttgaaaaaactcacaacacacatagttagaataatcaaaaaaaacaaactggaacagagttagggtaggtcaggaatg
 cataaataataatgttaattggctgtttatattatagtaggggttccagtcacacagaaggctatttagtaatttaagttat
 tgagtcacaaagtatacacagatttttgatgatggaggagtagtgcccccaacccccatgtgttctcaagggttagcta
 tatttgaattctgctggtagctggaattgggcaactaaattccagagatgtgggaacgctcgtatcaatgggagaaataaata
 aataaataataaataaataaagtatttggctggcctgggtgcagtggtcactcctaatcccaacactctgggaggccaa
 ggcaagcagatgacttgaggctcagggttcagagacagcctggccaaacatgggtgaaatccgctctcactaaaaatacaa
 aaatcagctgggctgggtgtgocacctgtaattccagctgcttgaggaggctgagggatgagaatttgcttgaacctggg
 agtcagaggttgagtgagccagatgacaccactgcactcagcctggggcgaagagatgggactgctcactcaacaa
 caacaaaaagcaattgttaaaagataaagaaaaacagccacagatggggaatatttgcaaaaaaataatctgatg

FIG. 7

43/157

acccaaaatacacaaagaactcttaaaacccaacactaagaaatcaacaacaccccatataaaatggacaaacggggccag
 gcatgggtggcggtggctcacatctgttaatcgccagcactttgggaggccgagcaggccagatcaccttaggtcaggagttc
 tcaactagcctggcccaactagtgaacccgtctctactgaaaaatacaaaaattaggccgcgggtgacgtgcatctgtg
 gtccagcgtacttgggtgtctgcggcaggaaaaatagctgaaacaggaggaggttgcaatgagctaaagatcgccgcactgca
 ctccagcctgggcaacacagtgagactcogtctcaaaaaaaaataatgggcaaaagatgtgaacagacacccatcatcaaa
 aagataacagatgggaagataagcatatgaaaaatagcttaacctgatgtcattaaaggaattacaattgaagcaacaag
 gtaccactaaacccctattaaaaatgggtcaaaatccagaatgctaaaaacacccaaatgctggtaggaatggggagtgacag
 gactctcattcgttgcctgggtggaaatgtaaaaatgggtatggccatttttaagacagtttgggcagtttcttacaacactaaa
 catagtcttaccatcaaatctaaacatttacctcctagatagttaccocaaatggaggtgaaaaacttatatccaggggccgag
 cacagtggtctacacctgtaatcccatcactttgggaggccaaagagggaaggatgtccttgaggccaggagttcttctctt
 ttattttattttattttattttttcttttgagacagagtttgcgtctgtcaccacaggctggagtgacgtggagtgatctca
 tctcactgcaatctctcgcctccgggtctcaaggcatctcctgtctcgggctctgagtagctgctgggattacaggtgca
 cgccacocatggccagctaatttttatatttttagtagagacagggtttcaccatgtcagctagactgggtcttgaaactcct
 gacctcaagtacactgcctgcctcggtctcccaagtgctgggattacagggtgtgagctactgtgctcagccaaaggccag
 gagttccagaccagtcctggccaacacagtgatactctgtctctacaaaaaaaatttttaattagtcagcttggtgg
 cacacacttgtagtgtagctactggggagggtgaggtcaggatcacttgagccacaggagtttgagggttgacgtgagc
 cagcatatgctctgcactccagcgtgggccaacagagtagagaaaagaaagagagagagagagagagagagagagga
 aggaaggaaggaagaaagaaacgaaagaaagaaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga
 agaaagaaagaaaggaagaaataaaaaaggaaagaaagaaagagagagagaaaccttatccatcaaaaaacccctgcgcaa
 gaatgtttatagctgcttttatctataatttgccaaaaactggaagcaactgaagaagctcccaatgggtgaaaggataaac
 aaactgtggttatctgtgcaatggagttatcattaaagtaataaaaagaaagacttagcaaaccccaaaaagtaatatgcat
 actaaatagcatattagtaattattaatgcattaagcaataaaaaagacgtgtcaggcccaaaaaagtaattatcata
 ttactaaaatgaaagaagccagctctgaaaaggctacgtgctgtgtgatttcaactatttgactttctggaaaaggcaaac
 tacagacagtaaaaagatcaagggttgccgggggtctctgggggacaggaaggatgaataggtggagcacgggaatttag
 tgcagtgatatttatctatatgatataatgggtggatccatgatgtctttgctcaaaaacccatgaaagtacacac
 aaagagtgatttcttaattgaactatgtgcttttagtcaacaatgtatgtatattagttcattaattttaacaatgtacca
 cactaaatccagatgttaataatggggaaactgcgtgtgaggggagggatcacgtgggaactctgtactacagctcaataa
 ttttataaatctaaaacttttttttaataaagctctcttagaaaacaataaaaaataaaatgaaagtcaatttttttt
 ttttttttgaggaggaatctgtctctgcacccaggctggagtgacgtggcacaatctctgctcactgcaacctctgcct
 ccttagttcaagcgaattctcctgcctctacctccagggttcaatacaattctccacctcagcctccagctagctgggac
 tacaggcatgcgccaccacgcctggctaatgtttgtatttttagtagagatgggggtttcgccatgttgccagggctggtc
 aagaactcgtgacctcaggtgatccggccggcctcagcccttcaaaagtctgggattacaggggtgaaccacatgctgggt
 ctaaaatgtccatttttaaaaggcagctctggctggggcagtggtctcagcctgttaatccagcaccatgggagaaagctg
 agcagggtgtatctcttgagccaggagttcgaggctgtagtgctgtatgagtgccactgcactccagcctgggtgac
 agagtgagactctgtctctaaaaataataaaaaatacaataaaattttaaaaggcaatctgcagcagagagaagatga
 agttagcaggagagcaggacagaggagagtgccatgaacaacaggagaggaaggaaggtcatgggtggcagctccagggctc
 ctggcatatgttcaattccctgttcccagcctcagggaagccagatgtcccccactgcccactacatacaaacctggat
 ccttgacatcaactctcacttctctcatgtgctctaatgagtttttgttacttgtggaactttaactaacatcattaa
 ccaagtggacactgtctcagagtgggtgttatgagaagacagcagccaaaagacagctgcagccaagcagctgggtcatg
 ctgtaactcctagcattttgggaggccgagggtgggtggatcacctgaggtcaggagttcgagaccagcatggccaacatgg
 cgaaacccctctccactaaaaataaaaaattatccgggtgtggtggcgagcgcctaatcccagctacttgggaggt
 tgaggagcagaatgtcctgagcccgaggggcagaggccgagtgagccgggtgctgcactcactccagcctgggtg
 aaagagcaaaactctgtctcaaaaaaaaataaaaaaagacagctgcaacaaatgtcaggttctgtgtgtttcttctct
 ttctctttttctattttaatttaatttttttagagtcagagcctccctatgtcaccaggctggagtcaggtggcagcag
 tcacagctcactgtagcctcaacctcctgggtcaggcgatcctccacctcagcctcctcctagctgggactacagggt
 ggtgtgccagacatctgggttgtgtgttttcttttttttttgagacggaggtctgtcctgcccaccagggtggaggt
 gcaagtggcgcatctgtcctgacacacactctgctcctgggtcacaagcaattctcctgctccgctccttgagtagct
 ggggaatcacgggcgcacaccacatgccactaattttgtatttttagtagagacgggggtttcaccatgttgggccagga
 tggctcgtactcctgaccttgtgtccactgactcggcatcccaagtgctgggattacagcgtgagcaactgcaccc
 ctgctggctgtcatgttttctgacatcatgtcaaaaggatactcatactaaatggcaacacattctcagccctctct
 tttctctctatctgctttaccacacacacatagctgctttaccggttttctctctcaaaaaactcaaaaaacccctcccc
 aacctctatccactctgtgtcctccctcaaacccactccattttgtattctgcatcataaaaaaccaagcagagttctg
 gggaggcagaagcctcctcttatgatattggaagggggtggatgaaattgagcacacaaggccagaatcctcttttgtgg

FIG. 7

44/157

aaagatgggggcagtgggcagaggggaagcagggtcatctcttccctctcttcccttcccttctcttccacaccacacgc
tccgcttgggtgagctcatcgctcctggtgtctcaatttccacctccaggagatcgactgccaaatttctaccgcgaatc
ccaaactccactctgagctccagaccatctactaattgctagtcacatttctctgtgggagcgaatcagataggtat
tatattctgctgtgactaacagagactgaaaattcagagtcctaaacaaaagacttctcttctgtgctgaaaagctc
ggaagtgcacagccctaggctggaatagggtactgctgtggggggtccgcctccttgcagggtgcccctctgctatctcc
agggtgtcctccccaacactgtccaagtggtgactgaaggaccagtcctccactgacatccaggcaacagaaat
gagaagaacagggtgaggaaggagggaaccacaggcgccctccagctgttccaaggaaaggtcctggaagatgctaaataa
caattccgtttatatctcattagccagggttagtcagctggcaacacctagaagcaagaaaggtcgggaataatctgtt
tctgtcgggttaccgtatgcatgactaaaaagagagcgctctgttcttaggaaaaatgagaaatagactgggttgggggagg
ataaccagcataggctacactgcagcccccaacctcctcatgcttgatgtggctgaaataaaaagagcatctccccta
acatttcccttcccttagtcttgttcccttctgttaccataccctagcagagggggcaattctgtgggggtgggtcact
ggcaaggggagggtctcacccaagctcagtcagggtgagcagagagaattgtacaaggaaagctgaatttccagctgggtct
ggaggaacactgaaaaccccaacagaaagatgtccagggtggagcaaatgcagctgagagtcttctgtccaaggggacaga
atgtccagcagataggacacaaggggcaggtctctactgaaagcaccaggggcaaggtcacccagcccttcaaccggctgg
ccaccaagcactcggctccactctgctcctctgccgtgccaccttgcacttttgtctgttaccatctcctccaccagg
agtgccttccgcctcctccatgtggttactctccagccacctcaggctcagatccagagccacctcctccaagaat
ctccactgtctgtgggaacccctcacatcacactgggagacgaagccaaagtgggggggtcagtcagaaaggtggtgga
ggtaggtccaatcctggctggactggcttgggtgccatttgggacagggtattggggagggttcagcaggaaggtggga
cctcaaggcttctcttcttgaaggggagggaacccagctgttgacagagcagcaacagcaggaggcaccacaggctgagc
cttgaacctggggccaccaggctcctgcagggaacatgggagcagctcagctctccagctcccatctccaggagggaaga
tgatgcagctcaggcaattggccacctggagggaagtccagctgggtggggcaggcagaggaattcctggaaggagctc
accctcagcactgggttagaggacactagagacagggttgggccaggagctgcgccacctcctatccccatgaataagc
agggtggccgggtgaccagctgggagccacctgggtgctcaggtaccatgctctgggcccagctcctgtgagctgtgg
cacaggcaagagcaccggcccaagggtctgcctggggcagtgacctgtctctcctaagctcagcttctctcatctgtag
aaagagggtgaagagtccttccccacatggttgtgtgagggttaactgaagtgctagggttaggggacccagctggggcac
cagcacacctgaggttcagtcacacaggccactgggtgaagaaaacctcgccacctccgcctcctctgaggcaggtg
gtggaggcacatgcctcctcatgtgtgtgcatgagtggtgtgggtgcagtgcatgtgtgtgcagtgcatgtgtgtgcat
gtgtgtgcatgcatgtgtgtgcgaatgtgtgtgtgtgtgcgcacatgcatgtgtgttaagctctggagaacaagcagcag
cttgggcaggagtcggggggcagggggggacatgcaaaagtctctccacagccctcagtgctaacaggctgaggtggg
ggatgacatgctttagaacaacctcaccctctgcctaaaactgctgtggtctgcacatcgggccaagctgtggg
tacttgaagaaggagcctgctcctgcttccaaaacccaggaaagtcttgaattccccatggaggggacctggggccagg
gactcccccaagcagacacagcaccacacaagggtctcaagggaccccttctccccacccgctcaggttagggatggaatg
tggacaagcagctcagggtcgggggtgcctagggtggaaggttcagaagcagacaggcaggagctgccaaaggggccgg
acatgggggttggaggctagggtctggaaaagagtttgggagggtcagccaaggggcagtcacagcctgggtctctctctgg
gtcatgtggggagccacggagagttaggagtggtgatgggatggccccagagtttgtgcctttagaaagtcacctctgtgtg
agatagagatggactggggggcagggtcagaggaagggtcatctcatgagaggggaaatccctgaccacagtcacaggga
aacacgggggtcggggagagggtgctcctggggaggacctaactcagaggctgatcttctcctcgtgctacccacccagg
agatgcaggaagctggtgaggtcagtggtggggccacacacctgctggagaggagcaggaggagacgggagggcagag
agggttagagggcagggtcctgaggacacctcatccatgccacctggacaattgaatcgagcttctcctagtgtggaat
tgggctcactgggtcatcttgcctatcttctcatctcagactggatcgggcccaagggtgggtggtggtggtggtggtg
atcagactgagggggggcggaatgtgtgggttgggggaaggaggagcttccaaacaggaaaggttggaggggcgctggc
tgaggacaggggatacgcccccatgagctctcaccctgctcagagctcgtgcgtggccctcagagacgagccatc
tttgagctggagtgtggggcagggaaacccagggtggaaggggaaaaagaaaaaggggcaggaatgagtgtaggggacag
acctgacagccacaggacatgcacagcagacacacacctccatccctcctggctcgtcagctcctcaaacagggtgacg
gcttcagggcccaacaggacccctgttacacacagacacaaagacacacagacacacacagcagtaatacag
aaacacagtcacacactagcgcacacacagacacagtcagacaaatccaagaaacacacacacacatgaacagacataa
gcgaagtcttacacatacaagacacacttaagcacagacacaaaacacacacacagatgacagacggacacacatagac
gacactggggggggcctgtggagactgctgatggataccatgttggcttggtatccagaattgttgggtgcagcgac
tccccatgctgaactcattaaaggagggcctcggggcagtggtcttgagaagaaaggcagcagcagctagggaagtga
ctgactgctggagcattttggctgggcagagagggaaaggagagcatcttcaggggagggaataagctcctgcgaagg
ttggagtgggaagacaaaggaggtatcccaagaaactgtgcctcatttgatggggctgtagggggcccttctctgagggcag
ccttccagggaacaaaggaggtgagctagaaagcatctggaacttctgtggggcacaatctgtggggcagaggggaaggga
caggcagaacgagtcagctgctttagaagctccacgaggacctaagcagaattcattcatctcgttcatctt

FIG. 7

45/157

atttatgacaggttctctctctgtccaccaagctggagtgccggtggtgtgatcacacttactgcagccttaaaactcctgg
gctcaagcaatcctccacccagcctcctgagtacctaggactacagacgtgagctactacacctggctaaatttttaaaa
cttttttagagacaggggtctccctttgttgccaggttggtctttgttgccagccagcttcaagccatcctcccaactca
gactcccaaaagcgctggatcacaggcatgagccacgggttccagcccagaaagagaccttcacccacaaaacccctgt
ctggaggccaagtcccagcatccccagaagggccccgggttaggacaagaaaaagggctgtgtgcctcggaggagaggctgg
gtgccccctggtgcggagtggtcccccttgccgggcatggggcctcagctctgtgagactccgcgcctcgtcgtctgaac
tgagcaaaagacagagctaggcctagcctggggaccgcaggggcctgcctcatccgcaattctgcctcttttctggccact
gaggacccgggctagggaagctgtaggggcagtaggggtgggaggagagagggttccgggtcaagggaactggagggggtctg
gggcagggaagtggagggtcttagggcgggcccttccccctccgctgttggggcctcgtgaaaggggcccagggtgagcc
ccaaggtgggctgtgggggcccggggcgggggagccctcagcagctggaccgcaggaggccgacagggtctggaactcc
tcttccccctcctccgcgcgcctctccctctcccccctgggctgctaggggagcgggggaggcgaggaggggc
tcaggggaggggcctggagcgcgggacccaggtctggccctcctggcgaggcccgcgaggcaggccctgcgcgtgctcgt
cacatctgggcccgggagcgcgggcccagccggcgggcgaggcgccgagccatccggctgggggagggggcccggcg
gggtgggcccggcccaagaaggcggaagcctcagggccggcgcaagcttgagggtctgcagcaggcgggcagccggc
ccacccagctggggggcccggggtcgcggcgggctggacggggtggagctgtcggggaggcgaggtgagttctggggcgg
gggctgcggggcccgcagtgaggaaaggcgaggaggttgcgcctcccgggtgtcggctgagaccccacaaaacccct
ccagactttgtctggtggggacagatctgcaagccccctctctgcagcgtgggtcggggctcgggaaggcacttggggcag
cgactcttggtctcccaactgggtctcggggggaccacctccccctgctcctcgacagctgctcagagggttaggtctggga
aaactcctgtgtcttaggggctggcaggacgtggttggggctggggcctgctccaggggagccaaactgagggacccggc
taggtctgaggtggaaaactcgtggagttgagttgatgtgtgtgttgggggggtgctcaggggtcccaggccctgtggtc
ccgaagcgcaacagagtaactgagtgccccaggtgcccagggtctgaggctgggaactggtcttccggggcagcttgaggtc
caggtggggtggggctcagggttccgcctaggggcagtgagccctatctgtgaggttagcgtgtgtgtatgagagagactgg
aagggagtgatcacgaaactagggaacctccggagagcagacgcagcgagggaatccaggccagaggggttaatgggggg
ctgtggcgagagcatgggggggggtctctgttcttgtaactggctgctccatgatgaggaaccaaggggtgcggcagg
atgtagtgaaagcaagatagagacagggaggggccatgggaggccattctgtcactcctcggccagcgaggatggg
gcgcacactcactcagccacactgtgtccactcttcttaagggtctcctccaggaacagagcttgggttgggt
gcccgtttctgatctgtgcttggggccctgggtggccagctctgtggcttgggaagaaacgcctgtggtcctcacttgcgcc
ctgacggagtcctgttccattacagggcctgggtgggtgctgagcagagagcagcaggcggaagaggcagggtcctcagtga
ggaggggagctggaatgacctgggtgggtctttgtctacacagggcatcgtgtggagccctgaaacccagaacactgcat
tcccatagtgtcctcggggacatgagcgccagctgaaacaccccccacaggaagagatcctccagggtatctgtcaggc
aaagggttgggcccctgcaagggtggctcagccctgtaatoocagcacttggggaggatgaagtgggaggatcgcttgagg
ccaggtggtcgaggctgagtgagccgtgatcgtgccactgcactcagcctaggcaacagagcaagaacctgtctcaaa
aaaaggagcgggagtgggggcagggttctgctcctaaggctattgttaagaccaggggatactcgggtaggggtctcgtggt
gtgagttcctgtcctggggccacacctctcctgggtcccagtgactaatcggccctgcctattgggttagaagttagggccag
gggtcagggtgaagctgatgactagaggtcaatatgaagccaagaggttaggggtcagtcctagggcagggttagggatca
ggctggggccagagtttggcatgtgcttctgtgagccaggaaatggagttcctttctgtgttaaatcaaaactgtcccactcaat
tactctcaaacccataggcaccaccagctgtaaattagtagggccctgaagcgggggtcagccctcccctcattttcagct
gtggactgtcccagacctagcttccgggagctgggcttgtcccgtgttctgtggtcagcactccttaagtgtccctg
gggtgtgtgtgctcgcttggcagtcagccacctcttgggttaacctccacctctcggggaatccatgggaatcctcctctt
aacacaattttttttttttttttttttgggagacaggggtctcactctgtcaccagggtggagtcgagtggtgtgatcatg
acttaactcagcctcaacctccaggctcaggcaattctccacctcagcgtcccaagtggtgggtcagccttaggcagctgca
ctaccacgtcagctaatatttgatttttctgttagagatgggggttttctgtcttgcccagggtggtctggaactcctgagc
tcaagcgttagacttcaaccttgaacctccaaagctgctagaattacaggcatgagccacacacccagccctcaacaaat
tgatttagttctcactcctccttcaataagaaactcatgatgactagaatgcccattctcagtaagaccaaaatatacaca
gaaaattcagaactcgcattcatgaaggaaaagaacactgctcaaacctgcccactccctgcttttatcaaatccttgca
tatctggactgtttaacataagccctcgtgttctcacccttctacaggaaggggcactgagccagagatgtggcatgat
gcactcagcctaggtgcccataataggctcttgccatccctgaaccccagcagatggggcactcagcatatcgcaaatga
gagacggttagataactcgggaacagcttccgtagctggggaagagtagtagggccatgggttggatgtggaactcaggtgaga
gcggaggactcagctcttttttttttttttcttgtagatgggggtctcactcttctcggccaggtcggagtgcaatggcgga
tcttggtcactgcaacctctgcttccagggttcaaacagattctcgtcctcagcctcctgagctgggaatcacggat
gtcgcacctacgcttggctaattttttgatttttagtagagacaagggtttcaccatgttgccagggtgggtctcgaat
cctgacctcaggtgatccactgcctcagcctctgaaagtgctgggattacaggagtagccacgggcccggccagag
tctttttttttctgggtggttctaggttggttctcctccagggtacaaatagcctgggtccagaaatccattcatctca

FIG. 7

46/157

[illegible]

FIG. 7

47/157

caagagggtcttgatggggcgctcctgcctgaatgctggctcagggtgtgtgtatgtgtgtgtgtgggggggtgggaatc
 agacttccaaatgtgtgggacacccocagttcaggactgggctatctctatgggtgtgggaggagggtgctccccctcctc
 gatcctcctcaagcagcagctgggocctgtctgtggcccttgtgggcatgggagctgcagagttccctcaagtccctgat
 agaaccacgtggcccaacttttaggggtggaggggagcaacccacccctcctagccagtttagctctgtagctccacctttga
 agggaggagggtctctctcctcctactcctccaggagctcaacctgcctatctccctgcctctccatctctggactg
 agtctcctgcccacacaatgctttctgtctctctccctcccaactgtgtgtctgaggggcccctagtgaggggacagt
 gagctcctggggaagcatgggtcatctgtctagggaacgttgccctctgacagcagaccctcagctcctgcccctgctc
 atgtgtgtgtcttctacaccccgagctcctcctccacagcagagagccctttgccttagccagcagaaggactgatggaga
 acctttgggtgctcctggcttggggcccccgtgcctacactacccctgcttgtgccaggatgaactgcccctctctctctgca
 ggctccagcccgagcggtccccagtgctagcccaacccggctcccccaggaccaggaaagctgcggcattgccccct
 cacacccctcgagctctccagtaagcccgagcagggaacaggtgggtgggtgacctggctgggtgaggacgggtcgtgct
 ggcaaatgcatgacagggcctcagctaggaaggaggaggggatgggggtagccacctgctctgtccccgtacactcca
 gtcatgtgccccccagaaggatggccttgggtgacctggactataaagggttaacaattctactgagaaggcagccccc
 atgttaactaactcccaacccccccacacaggagggcaacccacggccattgaagctggcctggctgaggcgaagtgtccct
 gagcatgtggctttcttaggtccccagagcactgggctcctgtgggacacagcccaacttaatacaagtcatctctgaa
 ggagaggagatgtgtgtccaaaaaataaataatagtctgggaacacttagctgggtgagggatgggtgctatcactc
 gtaatccccagcactttgggaggccgaggtgggcagatgtctttagctcaggagttcaagaacagcctgagcaaacacagc
 gaaaccccatctctacaaaaatacaacaattagctgggcgtgggtgtcacacactgtaatccagctactcgggaggct
 gaggggtggagaattgtctgaacctgggaagcggagggttgacgttagccaaagatcgcaacactgcatttcagcctagggtga
 cagagtggagacctgtctcaaataaaaagggaactactagcttgaagtatccaggacacccactcccccaccccaagctc
 cctgggtcacccacactggatgaacttgcctgggtgacccagcttagggaccccttatcttaggggagctccagccctg
 gcataggagaaggacacactctctggggacatgcactgcccctctcctctctgcccagagaatcagacagctctgtgtg
 tgtcactatgggcccacaaagagcagaaggatataatacactgggtccatgtgtctgtaaaactgctcggaggagcaaaa
 ttgctccagctgggcccacagctaaatggctgctatgaagatgggtgtgaggtgggggtgggggttaaggcagtcagagggg
 aggggtgggagaggctgggaagcaggaaagcaggagcttgagccaaacccctgcccggggctacccagagacacgtccaag
 ctccagctgggagcctgggagggcagaggagcccgagcaggtgggagctggggcaggtccgggaggtcagggctgttggga
 ccacagtcagtgaggggaggtgtcttccccctcagaggaagcattggggctgtgggggagtggggggtctccctgctggct
 gcttgcataatctcaagtcttgcacttgcagaaaccccgaggtccgagccccccagcaggctcctctctgtgggtgggtggc
 cattgacttcggcaccacgtctagtggctatgtcttcagcttggccagtgacctcaggccatccacatgatgaggtgag
 tgcggctgggctgagagagtgagggtggggagggtggggagttccctacacttgggtcccaaaagtactgtccacagagaca
 tgggggtctccttggaggccctgccaccccaactctgggctccccactgggttaaaagtggaggatgggccccaggcct
 gggtccttggcctcaactggagcagctcccaacccctctgtaggaaagtacagacttgttttttaattttgaattttttttg
 agatataacaaacataacaataaaatatacaaatgtttg
 tgcctaggctggagtgcaatggcagcatcttagctcattgcaacccctccacccctctcgggttcaaggcagattctcctgctc
 agcctcccaagtactgggaatacaggtgtgcccactatgctggctaaactttgtgggtttttttgtttttttttttttgag
 acggagctcactctgttgcaggctggagtgagtgacgcaactctcagctcactgcaagctctgctcctgggttctcat
 gccattctcctgctcagcctcttgagtactgggacacaggcgacccgcccacacccagcctaattttttttttttttttt
 tgtatttttttagtagacggggtttccocgtgttagccaggctgggtctcaaatccctgacttcgggtgataccgcccagcct
 cggcctccagagtgctaggattacagatgtgagccacacacccctggccttttttttttttttttttttttttttttttttt
 ctctgtcacacaagctggagtgagtgagtgatcatagctcactgcagcctcaaatcctcagctcaggaacactccccc
 actcagcctcctgagtagctaggactacagacagcaccacacacccctggctaatttaaaaaataaataaataaataaata
 ggaagagacagggtctcactgtgttgcagggtgggtctgaattccctggcctcaagtgatttoccacttttagcatccca
 aaaaactcgtggaattgcaggcgtgagcaccatcccgagccacaaaataattttctaaactgtgagcaaacacacata
 taaaaacatatatgtgttttttaagtaaaagtagatatcaaaactgtctgttttaaaagcttcaaacactattttttagt
 atatttcatgtcagcacaacgaagctaaccttttctgaaacggcctagatttccatagctacagatgtgtcattttta
 tccattccccctattgacagtttttaggggggttttttctcctctctatttttcaactatttaaaaaagtgacagcagtaaat
 cctctcttggaaacttgtgcgcaactccatggggtagattcctagagggtgattcctaggtccaacatcctaaactcttt
 tttcttcttt
 ctaagggtcagtgacaaagtccacagttgttttgaagaacatgtttcaaaaaactcaaatgttgggtttttatgtg
 ggacctcaaggactcgggacctcaagaccccgactcctctgaatatgtcctagagggtcggcagggcagtgagggtgaagctga
 caggagacctcaaagctctcaacccagactgggagggaggttaagaggggttcccaaggcctggagactggatgggaagg
 gtgaatggagacatagacacagcaggttaactgctccaaaatacatgggacttgagctcactgagactggactgctctcc
 tttccaatcacaggaaacccaaaggctctgagcacttcagccctcaggagctgacagcactccagatgctctgctg

FIG. 7

48/157

atcaacacacatggctgaggccctgctggggctaaaggacacatccagggaacactgtctctgtcctgtggaggtgacagtc
 acaggaaataactgggggacattccagcctctcccaagctctctctccaccagacacccaagccctaacatctctag
 aatacccatctacttctctccatggccactgactgctgcccaaggttcagcccgacctctctcaactggatggctgcaat
 atcctcctccttggtttctcttctccatgctcccttccctgcattctgtgtgctccatgggagccagtgaccttctcttct
 aaaggcaaatgtggctcatcctctctggctactttaaataccacgaagacagaattctgtaaggccctctctgtgctctcactg
 atctcaacttagcattctctgtgttttctcctgctgctcatggagccgacccatagctagttctcctcctgctcttccacatc
 tgtaaagtttttgcaatcgtgttcccttctgctggattccaccccaagaaggagggttgcactgactgctcctgctcctc
 cgcaggaaatgggagggcgagacccgggctgccccacagaagaccccgacctgctgctgctgactccggagggcg
 ctccacagcttggctacacccgcccggattactaccatgacctggaccccgagagggcgcggaactggctctacttcg
 gaaggttcaagatgaagatccacagcgcccggtgagtcacagggctccagacagggggcgggggccagcatggaaaagg
 gcagggtcaatgggggtgggtgggacaaaaacaaacgtgtgaggacccggcccgatggagtcgtggctgagagggggcg
 ggctaaaggggagcgtcggaactccgggtgtggggcgagctcagaaaatgaggtggagggcggggctaatgtgggtgggctaa
 tagtgaagctggggtgacaggaggggtggggtcaaggagaggggtcgggggcagagctaatgtcacatggggcaagagtg
 gacgggtggttaaaggaggggggaagctccaggaaacgggggtgatttgaagcgaggtcgtcaggaaatgagtgccaagct
 gaggcctcctgcagagcccgctgtgctcctgccaggatctcaactgaagacccagctcagggcagtaaatggaagac
 gatgcccgccctggaggtgttgcgccatgccctgcgcttcttcaggggagcacgcccctcaggtgcgctgcggcccaact
 ctgcgactgtggcgaggacccctatttccctcctcagaaacgcctcccccacccgctcccgacattggagattgggtga
 gccacccgcggagctcagaggtcatcttctccagtaacctcctccttttctgtctggtagagcctgcacaaagccatct
 gatgggagggggcgagttctccagctcctgctgggaagtccttccctgtgatttgattagtagtacctccagttccgcagagg
 gctgaagacacccctccctccaaagcagcttccctcactgcccctcctgtaccaggagctgagggagcagagcccat
 cgtctccagagagaagacactgtgcctcgggtgtgtgacggtgctgccatctgggaaacagccagcacaagcagttcatgcgg
 gaggctgctacctgggtgaggaogtgcaggggggcccgagaaacactgctcaggaaagggcgagggctgtcccatgcttgc
 atgcaacccacacaccccttgagaccacagagctatttgggaaagaaactcagcctgctcctcagggaggttggtagagttg
 ctccacacagaagaggaggtgggctgcagggaacaggggacagaggggacaaaagacacagcccgagtcagtagcaggt
 gccctaaagtccaggtgtcctaaaagcagagcccaagatggagagtccttttttttttttttttttttttttttttttttt
 gctcgtgcggccaggctggagcgagtcggctgactcgggctcactgcaagctcgtcctccaggttccacgccatctctcc
 cctcagcctcagcctccagtagctgggactacaggtgcggccgacacgcgggctaatttttgtatttttttagtagagac
 ggggttcatcgtgttagccaggtatgggttcaatctcctgaccttgtgactccgcccgcctcgggctcccaaggtgctggga
 ttacaggcgctgagccacccgcgcggccgaagatggagagctcttatatgtgaggtctattgaagaagggttctcaggagaa
 aggcgagggagggagggcgggacaggagagaggtatgaaggatgtggtcttctgctggagctctagcctcaggtttagcct
 cagctagatcccatggggagctgcagaggggagactgtagcagaggttggggcgggctctctgcacatccgtagcgtc
 tgtcattggttctggggaaatgggaggggaaacctctccaaagtggggccatcccatcagctgagggcgagttatccagag
 gaggtagcagctgtgagccaatagcagccaacactcacagtggaaggcaggggacccagaaacatcacttccacttggcat
 cagttattttgagggacacccccaccccgccacactttgttgggttcttgagcaggggaaataggactatgaaaataagg
 aggcgattctagccaatgcatacaggtcttctcctacactgagatgctcaaaggtctagagcaaggagttggggggggctc
 ccgtggagacagagagggtatagcctcaagctcatgtcttggtcgtaaacctagagccttggtacagacccctaagctg
 tccaaaagggtatctcaagcctgcctgtggacacccctggagggaccccaactcctccatccctatggccaggcagtgacc
 caggggccaggtctgtaaatgtctacacttttcaactggaaggagctaaacctgagcagctggacatccctgtgcaaaagat
 caggatggaaagcaaggctgtccccagccagaaacaaacactgctccccagcgctcccttgctctattttccacacctct
 tctcacgctgttccctggctgctcagggaaagtctggccctgggtaaaagtctccattccacaccttatcctctccctcta
 ctgggtcagaactctggggaacccctacaaaacataataacttcccaagggaatccctcagtaggctatttggcgaggtccca
 gatctcaagtggttttctccacccaactggcctgtcttggtgctgggaatttcagcttctcctcctgctgtatcagattcc
 ctgggaacaaagtctcctgggaagtggtgtggcctatggctacgctctgacacccagctcagtggaagaaagggttcaat
 gagctaaaagggtcccccacccctcctcagggcatcacacaccccaacttctgctgaggtctctcggcaatccccca
 gccagcgctcagctggcgctgtacctgtgaaggggaaaccttgcctgagcccccgtatcttggaatattgaaggggca
 gctccagggttctaggtctgctattctgctctgagctcttctcctgctgggacatcttctggcatatttggccaggggcca
 tccctcccgagcctggctcagcgtggctatgaggcagctccaaaatttctgacgacacagaggggctgagaggcctaac
 atgggtgggggtgtgactggaaggaggggtgggagcccccagggcggttccccacactcacagtgccatcttaggtgtga
 cagccacagctgctgctgacccctgcccacacccatccccaggctgggaactggtcccgagagaatgcagagcagctac
 tcatcgccctggagcccgaggcgccctcggtatatactgccgaagctgctgctgcaacagctcctggaactgagtgggcgg
 ccccaaggtgtggggcgctcaggtcagctgcgcctccatcgactccagcttccgtcaggtgagctgcccggggacac
 accacccctcgagggtcagagggtcactgaagccagaagctcagccatgtctagtagaaggggagagggtacccacc
 tggaggagcccaatttgagcaggcagaaactggctgggtggagcctgtctgaggaaggaggcacagccctgcccagggt

FIG. 7

49/157

caacctcgctctcagggtgggacatgcacctacctaggggagcccaatctcagggtggagatttcaaccttgccctaggggag
 caaagtctgaggaggcaggactccgccagacccctgagttctgagggttgggggagacacagactacccaggaggaaact
 gctgggggtactgacaaaggaggaaagtccagccactgcatacagctgtccctctcactccctcatcttccccgacacacat
 cagccccaagctcctgctgcaggagaccaggccacaggccaaaggacagctatgggtcatctctcccaaatgccatctccag
 ggcaggaggagaagccctaagtctctgagttccctctgagactccaaagacctaactccctccctccctcccaaacccctctaa
 cctgattttgcctgacctgagacctcctgcgttaaaggaaaggccctgtgtccataaatatttccccacagctgttggga
 tacagggtgggagtctgggttccaggttgcctctccagtcaggagcagggtggagtttcaggagcactggctgctcc
 cagtgcccatggagggtcctgggcaggaggatgggagttgaacgcctaagctggagcactcctctcaatctcactccctg
 ctgtctcctgacccccagctcgggagcagctgcgaagggtcccgccacagccgcagcttctctgggtggagtcaggcgttagga
 gagctgtgggcagagatgcaagcaggtaggggggaaagggggacggagtggtatccttgggccctacccgggacccatatac
 tgatggggggaagggtatgtttgcaaacccgctcctctcctcctccattcgtgtacccaaccccgctccctccacag
 tcaccgcacccccccccactcacagcggcgccctaaactcccaactcctccaggggatctcccgcgagcgtcgggtgg
 agttgcagagcctctgggaacatcttctccccacacccctgcgcccatactgggtgggtctgaggttcaagcactggaagcc
 cctcaagtcctcccccgacccctgcagacaggcttgggaacccggggcagggtggaggtcgggcaggctggagggggc
 gcagggtcagggtgcgagggcccgccacagctgtgtgcccgcgtcgcgcgcaggagacccgctacgtgggtggcgagct
 cctcagctcctcccccgacccctgcagacaggcttgggaacccggggcagggtggaggtcgggcaggctggagggggc
 gcaagggtcagggtgcgagggcccgccacagctgtgtgcccgcgtcgcgcgcaggagacccgctacgtgggtggcgagct
 cctcagcggagccagctgacactgacgtgcagcgtggagcagcccatggacccctcaaggagctctcaaggagctct
 ggtgagtagccaggcgcgcgcgcgggtaccagcgcgacccgggtcgcgcgcgcgacactgcccctcggcgcccgcgag
 cgtgacgcctcttgcgccccctgctccacccagggggccccttatggcgcggtggggcgtggacctggcccttcgagcagc
 tgctgtgcgcatacttgcggcaggagctcactgcgcaaccttcaaaaggcaacgggcggcagcctgggtagatctgacactc
 gccttcgaggtcgcgaagcgcactgctggccacacccgtgcaggggcgctcaacatctcgtcgcctctctcctctcatgga
 ctctcacccgaagcagcgggggccaacgctggagacgcgtcgtgcgaggagcaggtgggtcctggcccgcggggtcagg
 cagggttctgcgaccccggaatgacgctgcactggagggtcggcgggcccaaggaaacgctgggggtctcgtcgtgattcatc
 ccacataacactaaagccagcagggtcgggtggggcgggcgaggcgagtgagtgccccagccacagcagcagctcc
 accacaggaatccgcagccgaactggggcaagacagagaaatcatagcggggagcggcgaatgctctcctcccagcc
 tctctacacccccaaccccgggccctcggggcccatgctcctcgttctccctgcaccaaaggaggggagggccctccca
 ggaactcgttaactgggaactggagcaggctgggaactaaatcctctgagtgagtaggggtggagataaaggagactaaactcc
 cgaggtccagctcctccagacacacagtcagctcggtgcggcaggcactctcgtcctggaggcagaggttagagaaataaggac
 cagggaccccaactggggcaagcgtggggccctgacccgatggatatttgcctcttccacccaacagcagctgaacttc
 gtgaagtggctcctcacaggggatgctccgaatgtctgtgaagccatgaacgagctcttccagccacccgtcagcgggat
 catccagcacataggtgagcactgagcttggtccccacccgccccctacatgaacaaacagatgcagaataatcccccc
 ctatcagtgctagatacctccacacacccatacactgtgatgagactagaatcctatagaacacccgtcgggatgaagt
 gcagtggtgattaaagctagagggttgatagtgtcttgccaaaggcaaaaaactctcgtgctcagtttccccattt
 ataaaaatgggggtgatagttattgggttctcaaaacggttatcacagggtataaaatgagctgaagtacctcagagtgagcaaa
 tgtcttgacacacatgtctaggtgttataacgtgtgaaaaatgcataccttatctcgtcctccacgtcgtgggtgggag
 agaagtggggagcgtgagttgtggggaggcgaagccctcgaggactccgtgagctcctcaagaaagtgcctcaaatggct
 actttctagtcgcaggttaggtacaggctagggaaggggagcgccgttgcccgccctagtgtggctggccctcagtggtctctc
 tccccgcccctctcctctgcccctctcaccocgctccccccgtcctgtcccgcagaggccctgctggcacggcgagg
 gtgcagggtgtgaagctgctgtctcctagtgggcggtctgcgcagctcagcgggtgcagcacgggtgcaggcgccgt
 gggcgcccgcggtctcgtgctgtcgtggtcccgccagcagctgggctcaccatcctcaaaaggcgcggtgctgttcggccagg
 cgccggcggtgggtgcgggtccgcgctcacctatggcggtggggcgtgctcaacgctcttcttccaactggggcgccg
 ccgcccgaaaaagctgctggttcgcgaacggcccgctgggtgcaccgacgtcttcgagcgctctgtggcgccgagcagtc
 ggtggccctgggcagaggaggtgcggcgagctactgcccggcgcggtcccggccagcgccgctactcatacactgactact
 gctgcggcgagaggatgcgcgcttctcaccgcaacccggcggtgcgcaaatgtggcgctcagcctcagcttgagcct
 gcgactgcggccaggaacacggcgggcgccctccggcgcccgcgagatccggcgcccgcatgagtttgggcgacacga
 aattaaagtcacccgctcagcgtcagcagcaactcgtcctcgtgcgcgctgctcaacgctcttcttccaactggggcgccg
 ccggcgcggtgcagcgccgtctgcggcgcccgccctcttgggttcaaggggctgcggagcggggtggggcgggggaa
 acgtagttctcagctctgcgccttccaacccctccagcccgggggagataaggtcagtgaggagtggtgggggacac
 acccagagactggcttggggatggggcactggtcgctgactgcaggctgaagggaacggccagggactgaacgggttaa
 gagaagaggttgcgaagacagagcgcgcagccggcgaagggtcgtgacccccgaaggaaagaaacgcaacaggaagctcct
 ggtctgagctggcagctcaggggtgggggtggggatggcaggaggagccgagagaggttggtcagggtctggga
 cctgcagggtcgaagtcaactcatcgaccgactcagccccaacccgggagccaggcagaaaaacccctgtgcgtaggaaag
 tgactggaagtggaactcagagggaacaggtgtggtggcacagctcctggtgtggtgtgacaccccaaatgactgtgaa
 ttgtggaaaggcagtagatcctaatgtggagggtgggaacattatgtggtggaggaactatcagggttagactctt

FIG. 7

50/157

gagagacaaaacaccggtctgggaagggcccaaggctcagcttatgaaggaccccaacttgacccccaccagccatggaa
tagcagcattgggggttggttagtgggggagccagaggaggaactgaggggtgtcccgagcttgcctatgtgctcgggtatgctt
taggcacattcccttccctgcactctagggctcagtttccctaaagtgtgaattgcaggcacttccctctaaactggcatgcac
cagccccccactctgctgagagccctgaggtgagacaataaacatttatgtcataggggaagccacagcctgtgcatatgggg
tggagacccctaatagtgaggagaaatgcaggaggttccccgtgtcagagaaagggagaaagtctcagctgtgcatagg
gaactgacagaaagggggtgctgctgtcttcccatcaagcatcccaaaccaactcactgcttaagaacctctctggcctaca
catgaggttccctctctctcattcaaattaatgtcttggaaaggcagctcttggcctaaaatgccaccctctgtcatc
ctcttgggggttaggtgtataataacagcgggtgccctgctcttgagtgagctaccceagttcttccctgggga
tctgcaaaagggttaggcatatcccacccagagaaacaaatccactggcctcgggtatcactcagggaagtattt
tttaggggtctatcccaaccagctgcttaaaacccaagttaacacagacgtcagggggtgggggtcggggagctgaacctcc
tctcatctcatggtggacagcagtggggactagggaaggggcaggagaggtgctgaagcaagggcagcagtaatggggccac
gacgcacagagccagctccctctctccagagcccttgggggagtcctgtgcttgggtgggggagtgggggagccac
ccaggccctccctctcccttctcagacagcctcttggggctcaaccatttctccgggaggagactgaggcaaac
aaggagagggaagtgggagagaggaagagaggggagggggcagggcagcacaacatgaaggcagaggtgcaggagggtcgggc
agggcaactcacccccagccacccccagagagggggcagggaagccacacatcaagcagcatctgaggggggacaggg
gggggttaaggcttagggggcccgggggcaggcggggcctcgggcctcagtcacaaagccgtgcagctgctgtctctgagtc
gtatccagctcgggccgacacacttgacacatccagcagcatgggcgtgcccgtgggtgggtctgcaaggcagggtg
caagtcagtgccactgtgcccggccggcccccactggggccacttaaggggcacccctccctctcagggatcagct
ggagttaggagctgccaggaggttgtagaacctctgacccgggcaggagctcacttgtcagcttggttctcctgaggag
gagcaaaaacctcgactcgagcttgtgcatccctcctcagcccaaggagcctccagctctcaacccatgacgcgggctgc
accgtcacgcgcacacaggtgccagtgccaccttcgcaggctagcagttatctcctctttaggacagccctcttctgtcgc
cgagttcacctcaccttgacactataacctgccccggggacaggggcaattggtcagcacctcccacgctcccactctg
ccctgggcatcaaggggcccttaccacacacatggggggactgcattgaagcccccagaagaggggacgggattggc
cctcttgaggcgagacacacatccccctgagcaggggtgcaagggtcaaaaacctctcagtgagctgggacatg
ctccagcgtctggggccaccccccaatgtgtctccaaaggtgaatgaagttccaaatccaggtctccaaaagaaaggtct
ggctgttttctccaaaagggaaggcaggggagagggcggtgatgaggagtgagggggggcaggggcagggttaggcttgagca
atccgatggccaagaggtaaagggcctgagggggtgtcaatcaatgaaagggcaaggtcatgaaggagctggtgacggggaca
tgaactgaagctggagctccacgaatactgacacctgcctctgcagggtggaggtgggggtcggggcagggcctgaa
gtgagctgataggacagcagcactgaacccaagggtctgggaactggtgggtgctaggagttatccacaggggtgtgta
gtccacagcaggttttctcagagggtttagggaggaggtaggggaaggtggagggtctcaaacagttctccagcagctcca
tcttggttactgccccaggggaggttaacctcacccatgggcaaggtacagggaagttatgagctcagtgactctgttcag
gacaggggacagaggctggagttggggacctgccactgctgccacagtgcaacctcacaagagggtgccacctcccgact
tgtgagctgccatccacctgacttggaaagttaggggaccccccatgtctactccccacacctcaggccctctgtg
gaggggggacacactctcagggaagggaacacgttggaaggttgagggtgcagggtggggagcagtgactgcggcagca
ccttgtcagctgtgcatctgggggcagggtgaaggttaatctbatcctgtcaggagttctcagggaagccaactcag
caggagaggtgagggaggaggacaccttcaactctcgtctccacaggcccccaggttgtgccacctatgatcccatgc
gcccagccagctgcttcccagagagccagaaagcccaactcctccatcactccagctgatgcaccccagcccttcc
acagacaccttgagggttcagcctcagcttggtgccacagagctcactcagcatctgtctcaaggggccccaggtcttc
gtggctcattctatagacaggaaattgaaacaaaactgactggccaaagctatgtagtggggcagggtcagctgc
tctctagacttcttgccattctctgctatgcatggcaacacagcaaggccacccactgctctacagggaagtcaaa
tccatgggtcattcacctgtccccactcagtcgagttaccagggtcagggatgctgggtctccattcgctcataggt
gacacaggccacagatagagctgggttagggatggggtcaaatatagagattagcagtagacagctgtgtcagccag
agctggcccacagacttggcacagcactgaactgggggtcctgtgctcctgtgccactggaactgggggtgagctgc
tgcccactcaaaagtgctgttccatggggccatctcagcaaccaagcgaattctgtgttcattcaacagtgcttggtgac
aacacccctctgggacaaatgccattcttggaaactcagtagtttaggggtcacaaggccagggtggtgtcagcaggg
gggggtggtgggggtgccccggcaggcaggcaaaagggtgctcgtctactccagcttgcatgggcaagcagagttcgtc
cacagttctggggtggggcagcccatctctggggggccaactgggaactggctgcagagagttttagaccttcaatggga
ccagggtgatattgctacatgacaaactcaatccatgtgccocctcagctctgctgtgggttagggggacacatccctcag
aaggcttctgtgctgacctgtaactgtgacccctgacccagggtactactatagaatttccagggggaaatagag
atgtgcagaaaatactctctgctgccatttataacatgaaacaaaacagcgcctaaccaacgggaacaggagcc
agctaaagtcaatgatgacacatcatggaccagtagggcattgtataaatctcgtggaaaattattctgacacaggtga
ggagacagttcacacaggaagcaagcaaaaacaggtggacacaggtgagggcagcttggtttctggaacaaacaggtt
ttaaataagctggatcactcacatcagtagtgttttcagaaaacaaataaataataataacattttaaagactgaaat

FIG. 7

51/157

aagccagggttggtggctcagcgctgttaatccagcactttgggaggccgacacggagcggtcacgaggtcaggagatcg
agaccatcctggcccaacagagtggaaccccgctctctactaaaaatataaaaaattagctgagagtggtggcagcgccgtg
taatccagctactctggaggtcgtgagggcaggagaatgtctgaaacctgggagctggaggtgtgagtgagctgagatggcg
ccactgtactccagcctgggtgacagcaagactcgtctcaaaaagaaaaaaagctgaaataacacacatcagaatat
taaccttggttatatttgggtagtattgactgacogctacgctgaagaggatgogtattgactgtgctttccctctcttgctg
gtggatataaaaaactggaatggacattgtctatgacgggagtgaggtctcttacaaggctgagcatgagctaccataca
atctagcaatttcaactcctagatgactacccaagtgaaagtgaaccttatatccaggggccgagcacagtggtcagcgct
gtaatccagcactttgggaggccgacacggagcggtcacgaggtcaggagatcgagaccatcctggccaaacagagtgaa
accocctgtctactaaaaatataaaaaattagctgagagtggtggcagcgccgtgtaatccagcctactctggaggtgag
ggcaggagaaatgtctgaaacctgggagctggaggttgagtgagctgagatggcgccactgtactccagcctgggtgacag
caagactcgtctcaaaaagaaaaaaagctgaaataacacacatcagaatatataccatgggtattttgggtagta
tgactgtgggtacttttaattttctctatatctctctgtcttccaaattttctcatgtctatgaagacatgagtggtatt
ttgcaatgaggggagagagctatttttaagtgttggtttgttccatttactctcttggggaagctgtcagctgtaggaca
aagagtggtgaactcagagttgggacagaatgggatacaaaacctgggccccagggtctctttttgtttgtttgtttgaga
cagagctctttctgtgccccaggctggagtgagtgaggcaaatctcagctcactgcaacctcagcctcctgggttcaagca
attcttctgctcagcctcccaagtgtctgggactacagcgctgcccacacagctcagctaaattttgtatttttagta
gagacgggtttcgacccatggcaggtgggacagaactgggccccagggtctctttttgtttgtttgtttgtttgaga
tcaatgggcccagactcccccgacccaagggtcactcctgttcaactctgggacccctgatatttgacccaactcagggcc
tcagctctccattttcccaatccagctccacccaagcagaaaaagggttgataataactctcaaaagacctaaactgag
ccagagatttggcagaagggtttaaaaaaaagtgtgacactgtttatggccatgttagctttcacacctgtatcacata
catacattcttccattctcacaacactctaggggatgtgattgtctccacttgagagaaagaaactcaggtgacctctc
caaggtcacagagccagaatgggtggcctagacctgaactcaggccactggcaccacagccacgacactgccttccatt
gtctattgtctgggtgaagcagatgacagcagtggtctgtctcatcttgactgtgagcatctcaaatgtctcatctca
ttgaaattcactgcctccactctgacactctgacacccaaggcaacgaaggcagagacagacatgcctataagccccggc
tgagaggaaacggggtatgggacccaagagacagggcagagccggggaggggaggaagctgctgatggggtaggttgcatg
taaccttgacagaagctgctgctcactctcctgggtgacaaatgaccacagagtcagagctctcactcaaaagtgcagct
ggctgtgggtactctctgcatcgtggcctgcggcaaggcggaactcgggactctgggtctgttgccatggaagaggtggg
agagggctctgaggggtttccctgtccccaggaaactcctccccctgtccccctgtcaccccccaagctgccccctgaca
tcatccagctccccctggctggctcttttcagggtctagatggacactaggatcatgaggtctgagggcctcccccgggc
ccccacctgccccctccccaaaaatccactccccctgagagtgagcgtggggggactcaggggactggctaatatgacccctg
ctttggaggggtggggctgggtgcccagagccaggaagggacagctctccagactcccccaaggccagggcagcgtgggact
cagccccagctcctctggatccactgcccagctgctccttccacagtgcgaagctccccctcctggggcctacccccaaagc
tgtgtcagtgcatgaggccctgctgccccactctccatgacccaaactgaaaggccagggaacaggctctcctctctgt
ctagcccccatcaaggccctccactgtctctgagctgggtgggtttctccaggccagaaagggcagacggatgagaca
ggcagggttagggggagccagagctgaaggcctggactggatgaatgtcaggtctgagaggccagagacagggccctagt
cctctctgacccccggccccctaggcccccccaacgaagccctgccccagggtcctcttggcgggaggaacctgatccagct
ttgagccggcctcttggaactcggaggccagggggacagctgggacatccacaatccccgttgagccccacagaaccc
ctcaaatctccacagatccacctggcctcaaaaacctccttatacctcttggaactctgaaagcaacccccaaactcctcc
acagcctgtctggccctcccatcactcagcagctcagtgctcaccataaggccctcggctcccccaacccccacagtgga
acggcagggacacagagcacacacactccacagacagacccataggagtcagcagggctctgacggggtcctgctagggtc
tgcttgagccccagaaactcgggacaaacactgacatagtcoccatcacatcgcaacacagacacagcactccacat
gcacacacacacacaggtggtgcgggccccctacctgcagcaatggccctcctaccaggggccacccggaagagcttgtgca
gtcctaccagccagggaaggcacttacccttgttggtgcagccatggagccctcctggacagcagctcctgcgcagac
acggccactcagcctcagcagctccactcctcctcaggggacacccagcgggagcctgaaaggatagggaagggtgggagggcag
ggtcaggggccacccatgggacccatggctcagggaacacagagagagctcctcttaggaaggactcaaacccctacgcctc
cagcagccccagtggtgctgtactctcagcaagggtgtctgtccccctctaccagctcctccagctgcccacttccacagag
ccctcaccctcctcctgataccccagtgactcaggggtcctgccagcagctgtgctgagctgagctcttaggttgccaaagca
cagcataggactcctatccccagcccccacagagctggtggcaggggagggcagctctttaggtatttt
ccagggaacagctctggtgctgctgacccccaaaggctgggccccggagctaccttattcaggtcacacaagcaatgcagct
gggtgcggtcagaggcaacagggtgctcagatgggtgggcaagaacaggggaccacataaagtacgcctgaccttagcctt
tccctcctgggtgatcggtcagggcccatatgcaaaaacagatgcctgtggagggggtgtgccccactgacccctcag

FIG. 7

52/157

accctgcccctgagctctcaggcctgcctccgcctcagctccccaaaggcagcctcagcacaggggcagtgagaggcaccoc
 aggaagccctccaatggggcagagctgggaccatctgagcatcaaaaagaaataatgagtgcaactgatggaaacagatggaa
 cacataaaagcagtgcaattcacagagattatagaaaaagagagaaagcaaaaacaaacccaggccctcaaaagc
 cctcatttgtcaccactgacggtagcagtgccctcttactctgaaagctgaggattaaagaggaaatccagcctgtc
 tcttgcccttttttgttaaccaaattgcctgggtgggtgatggaaagctctcttccgggaagaattcttgctaatataa
 acaaaagaaatgaccaaattgctaaagtcatcttgcaacccctaaaggaaacaaatgctggaggcaacaagaaccagtgatgc
 taaaacagagggggaaggttgacaggaagcagatactcacaggcgcccaagcgtcactgacagaggacctggtaggtttt
 tgtttttttttttttctgagacggagtttcaactctgtcacccaggctgggtagcaatggcagcatctcggtcaactgc
 aacatctgcctcctgggttcgagcactctccgcctcagcatcccgagtagctgggactacaggcgcccgccaccacgc
 ctagctaattttgtatttttagtagagacagggtttccaccatgttgccaggctgggtctcgaactcctgacctcaagca
 atctgctcacttcggcctcccaagtgttggttatcagggttgagccaccatgcctggccaggagacctgggttttaaga
 ggaatggcacaagtgtgcaatggaggtagtgactctccagtcacacactggtagctctactctctgtaggaca
 gctgtcacactgatgcctcctgaagtgaaggcagcccaagtcacagcatcactgacaagggaattcctgccaacaaaggtt
 taactgggaactaaaaccagcctctaacttcaactctctatttgctaaaaacacagaggagagggaacaaattaatga
 ctccatgagaaagcaatgagacgagccagaaaggcaggatttctacaggatgactgacctgtttttgtttttgttttt
 gttttcaaatgtcaagaagaaaaaagcaggctgggaacgggtggctcacgcctgtaatcccgacatttcacggactca
 caagggtcagggttcaagagccagcctggcccaacatagtgaaacccagctctctaccaaaagtaacaaagattagccgggtgt
 ggtggcgggcagctgtagtcaggctagtgtgggaggttgaggcaggagaatcactgaaatccagaaggcgagggttgcaag
 tgagccaaagtacygccacggcacttcagcctgggtgacgacgagagactctgtctcaagaaaaaaaagaaagaaag
 aaaaataaaaacgaggagatttatctagataggaaagcagcaaacactacagctgatcacacagctgatggttaaggttgctg
 cctgtgtgtttttgtttttgtttttgagacagggtctcactctgtccagccaggctggagtcagctcactgcccacg
 taatctccaggctcaggaatcctcccacccagcctgccaagttagctgagaccacagccacgcacaccatgcccgcg
 taatttttttagagacagagtgcgccaggctgggtctcaaaactccggagcttaagcgatccaccacactcagcctccaaa
 gtgctgggattcacagcgtgagcccgccaggctggcaataaaattttgcaaataaagttttactaatagacacacacac
 ttacttcaactacatgttctgtgtttttgcaataaaaacaaagttaactccaagtagttatgatagagacatattggcctg
 caaagtctaaaacatttatatttggcctctcagagcaaaagtttgccaacctgttcttagatcaacacacacttaaca
 gaccaaaaaaggataatcttcaggacaactgacctgttttttccaccacatcagctgcaagaagaaaaataaaggagagg
 gtgtttcgaactacatcaaaagaagacattttgggggggtgcggggcgctgggctgcaactgtaatcccgacactttg
 ggaagccaaaggcggttagatcacctgagtcaggagtttgggaccagcctggcccaactggcggaaccccatctctacta
 aaaaatacaaaaaatcagctaggcatggtggtgtgcacctgtaatccagctactcaggaggctgaggcaggagaatcgctt
 gaacctggggaggcgagggttgagtgagcgaagatcccgccactgcaactcagcctgggtgacagagcgagactctgtct
 caaaaaaaggg
 atocactgaggtcgagggttcaagaaacagcctggtctctactcaaaaaatacaaaattagcc
 aggcgtggtggcgcatgctgtaatccagctactcgggaggctgaggcaggagaatcgctgaaacccaggaggcgagggt
 tgcggtgagcogagattgcggccatgcactctagctggggcaaaaggagcaaaactccgtctcaaaaaaagggggggg
 agacattttgggggacttggaataatggcctggatagaggttaactaagaaatgagtgatcatgggaattgta
 tatgtaaagaattatgtccttattttgtagggaacacatactgaacaacttagagatgagtgacactggctcatcagtc
 agggacagaagaggaaactacagggtggcctcatagtttagggctggggcctgtggaggggcatgctcagggtcctctg
 tgccaagctgctctgggagcagatgtggcctcactcgtcccaactctcactgtgggcaaaagtcacccaggcccaag
 cctctgctcatggggcagggtgttcacagtttggttaacccaggagccccctttggccaatggaccaggccagccccc
 ctcttgagggtgagctctcaggccacagcgcagggtgcaggagtgaggactacacactcactcctctggccaggccact
 gtgggggtcagttgcctcagccctggaaagctcagcttgcctccaggggaaacttggtcttgaggagcagcctgcgttaacy
 ccccccagctacagctcccccaggctgtgaaactcagagtatatgacaaaaggcgaagcagaaggtgtgaggccctctc
 tatctcagaggctgctccagccacacactctcagccaggaggcagtgacagccagtgctctcagggtcccccatct
 ccaggatgcttctgaaggctcaggttaactgcaccagtggttgagcaaaaggctgcaggccttttcaaaaaaca
 ctctgaactggctcgaactcaggggcaacccctcctctcccaactactgctgactcagcgtccccccgcctgcttgc
 ccagggtcatccctctcagccaggaaacaaatccogtccctccagccacacacttgcgaaggggtctcagcccaacattt
 ttaaggatgcacccctgcaaatccagggaaggaactacagctcttttggaatccctttctctgctctgggtatct
 tgatctccctctctcacttacaagggttagtgaactgtgtttttgcaaacctctgtgccccgggaactcagccagcagc
 cagcaggaggagcgtgtgtttgttttgtttgttgatgaatgacacacacaaatctgctggagggtctggcgaagg
 cccgactctcagccctaaatgatatggtcaggggcagtgactgcaactgggggctctgagggaagcagaactccccacccca
 ctcttttgctcaccctgocctgggggggcaattccctctcctggccccacacagagtggtccactgggactcctca
 ggccttgctgagtcagctgggttaacctgcatgocctgtttgagagggaactgcccactggccactgagggtaggaooc

FIG. 7

53/157

acatctcatgcagtcgagtcactcagtaaaaggccaaacactgaggggcctgtaaccctctggatagtgacaacatagaggc
 aggaagcaagggaacttcaggaaacccaaggaaactgggaaacccaacactcctttctcaatggagaacctgctggctgct
 gtccctcggggagctagactctgtgcacacacaattcctcatcagaacaggcgaagaggggcactgggctccctgctg
 cagctgtgtccacagagaacagctaggccaccacacaccccaagcctggccctgctccagcctcctcctgtgccaaac
 aaggcctcaccatggctcaatccacatggccccagaggcagcagctcctgggtcaaatcaggatctgcccctgctagct
 gcatgggtgagttcttctcacttctctgctccctgttctcaatgtgtaacacaggaataggcagtggtcctctctccag
 gggttgctgcaactgctgctgaagctctgtgacaatgcaccccccatctctgcaaaagcaaacccccaaaggcctgggtt
 ctaagcaacagcagtttcagcagcacctgtagagggcacttcaggccaatcctggaggagccaggagtgaccttaga
 gtggggcccccagggaactcagccttcttggaaaaacagctcaagggggtgagggggcctccctcctcctgctccccccttctc
 ccatcccaaaagcagccaggtccctagggggctcagagaacagatgctggggagtttccagtcctccctaaacagaggggg
 tcacaagggaagtgtgcgaatgaacatcctgggaaactgggaaatgactaggggagaaatgggtgctcccccacgca
 aaaaaaattatacccttccccatgagatggagtgtagcaagctccaggcccccagcccagggtgcaggaaaggaggacta
 caccattcacctcctggtagcactgctcccaagctctgagacctcattctcattctctcagtcctccattctcattg
 tgggtcgaggtcttctctcttgagccagggaactcgcgaaccttccacctccctacctcttctcctacccccagcctg
 gggctcagtcctggctcagcactcagtcagtcagcagaagcactgtgtagcctccattaaagctcagcctgtgaaaaagaa
 caccattgagccttgagctggggccacatgcccgtgactctcaggactggacacagcagagggccacacatctca
 gaacaaagcctggaaggcagggtggaggtcagtagttgtggcagcttcaatcaactcagctttaatgtgatttaatt
 tctctctccctccagtcgggcaaaagggtgcaagataagtaggctgttctctctctctctaaacagtgaggtgctgggggt
 ggggggtgggggaataggagaagggaacctcaccaccacaccttctgctcccccacaagtgctgctcctctgccc
 agtacttctccacttctgcccctcagctagtgggtgcttagcctccagatagcatgcccacctaggccctgcccggggc
 tgtgactcagaggtcccaagaagcagagggcaggtggatccagggggtcagccaaagtgaggggtggagacacagagat
 tatctccagggacaggggtcgtgctgctcgttagctcaggtaggatagaatgtgggggatattccagctacatttctccctc
 caaaagccagaatgggagggggtaggggtgctgcccgaacttcttcaactccccgggagcagaaaaatggcctacctc
 cactttcagtgccaaagattcaagaagaagggaagcctcaccaccacaccttctgctcccccacaagtgctgctcctctgccc
 ggtccagaagctaaatccagtcctcctcctcctcattaccaccgggaggagggtgcagcccaagcccccaggcccccaggg
 tccaggtgtaggacctttatctctcctcgggagccactcctgtgggtgtggcaccagcccaaccccattctgtcactc
 tcagggggagggggaaatgtaatcggacatccccccatccaatccatcctgagctcgcaggggcgggctctgtcctc
 ggaataatcctgtgcactccccaccttcaactcctcggctgacgcaggaggtctccggagcccgactccagacactca
 ctgcccctccttctcagggtgctgaggctcggaggtcagagatgctactggtccaaggtcatgcagcaggagcagcgcca
 atgaacagggtcgtggggaggaggggggcgctgaccccatagcgcctccgctcactaccgcaactcatgcccccggaaca
 tgccttcgcggaaaaacccccagctgccaccagttaacggtcactgcgccccgggaacctgacactcactctagttcagggt
 gctgggggagctctccaggcccgcccccacaacagcggagccccccaccccagcctgccaatcacggccgctgggggtccca
 ggcaactgactgctcaggacagaaacgggacccgggaggggggacctcggcgaaocgggaagggaacccgggaggcaggagcag
 aggggtggcctcacttgcgggggtgcaccccggggcccggggaggggcagggaacgacacactgcagcggcgggcggtcagga
 gctcaacgcccagcagcaggaaggaggcccgccgcccggcgccctccgctcggcaacgcccccgctccgcccatttggt
 tgatctggagggtggggcgaggggagcgtccggaccaatgagcgggtccaaagaacggccaactggcgaggggcgccctc
 cgtcagctgccagggtgcggagggcagcgcctgctagtccgcgctgcggggcgagctctcgcgagggaagcggggcagg
 cgccccaactaggccaggggccagaaacggacactcgaagagggaaggaggggcctcggaataggccocgcccccgctcc
 tctctccgctgggggatagcgctctagccttgaaccttgccttaggacgcacctcccttggggcctctcgctctcgggag
 ggcgtcggcgcgctctcggggctgggtgagctccgaaagtggtccttcttgcgggtctcccgcggtctccggcctcggca
 tcgatacgggcggtgtggtctcgttcaggagctgaggaacctccatcactcctgttccgacccagggttgggaacctc
 tcccccttcaacccccctcctgtcttgaagaagcaacccccctgcccggcccgagagcgcgtccccgggtgctggccggg
 cctggagagcactcagaaacaaagcgggagcggggtgggggtgggagagctgtgagcgcccccggggagcagacc
 ctgctcccggtctccatggaggcgtcgccactggcagcgagatcgccctcgtcagcaccggggggggatgtggggc
 cctgcaatgagccgaggaagcagagggtggcctccggtctgcgggggttctacatctgagagcttttcaactctgaact
 cagcccatccactcctgccactctacccactcccccacactacactttagacctgtgctctgaaagtctaaagtgggac
 acggtggggcccaaggtggccgttaggtgattttcttagctcacatgggggtttttaaagaagtccctccccctcctg
 cctcctcgcctctcttggaatcaggataatcgtgttcgtggctttgaaatatcagaagatctgacaacgtgggtcacc

FIG. 7

54/157

ttccagcctggcattacaagccctttaagagcacataaatttaccatagtcocccgctaatacctgtttctcagcgtta
gttaccttctgggtctgtgcagacatttgtgatccctcccccctcatatccacacttccaaactggaaggccctccccc
ctctcttctcatagccaaaattcaagccctcttcaaccaagataaagtttgcctatttatttatttatttatttatttattt
cagagctctcactctgtgcgccagcgttgagttcagtggtgccatctcggtgactacagcctccacccctcccccgcg
ggttcaagcctctctcctgctcagcctccaaagtagctgggattctaggtgcgcgccacacacccagggctaattttt
gtatttttagtggagacagggtttaccatggttggtcaggtgggtctcaaaactcctagcctcaagtgactcgtcgtc
agcctcccaagtgctgggatcacaggcatgagccaccgcctggcccaagcctgcttttcaaaagttcttctagcc
aacttctcaaaaccatctcgtgtgtgggtctctcaaacacaaactggctgctaaactcagccagccgggttatttccga
attgtcaactcctttcagcaaacatttactgaatgtatctgatgagtaagttattgtgctaggctctgtggagtggaagg
cactcaagtttagttaggagggaacacacacaaagaactgtggtctcctgtccttgccaaatgttccctcacactcc
caggcatcgggcagctctggccactgggaggagctgaaatcacagagctgcccatgcttaggtaactacatgctcctcagg
gaaggccaaagatgggtctggaaagcctccaggagaaaaattagactcttgtgtcaggagggaagtccaggccaggagtac
ctggctgccccatcactgcccactgtgtctgtgttaccacacagtagtgccacggctctggggcttctcctgtcggagg
ccttaccaagatagatgggtctgcaagggaacacacacagggccctaccaggccctccggagacagaggcagggaacaaga
aatcaggaaacactccaggttaggagaacagatggagcactataaaacatccactcacggccggcggtgtgtgtcactg
cctgtaatccagcactttggggaggctgagatggggcggatcatgaggtcaggagatcaagaccatcctggctaacacgg
gaaaccccgctctactaaaaatacaaaaaatagccgggcatggtggctgatgctgtaatacccaactactccggaagc
tgaggcaggagaatggcgtgaaccccgggaggcggagcttgagtgagcccgagattgcccactgcaactcagcttgg
ggactgagccagactccatctccaaaaaaaataaaaaaatccactccactgagaacctggtagcttttaaaaaaa
aacaaaaaacacaaactgtgtcaactccaaacacacgggtttttgttttttttttttttttttttttttttttttt
actatgttgcggcggctgtgtctcaaatcttggtcctcaagcgatcccccgacttgggctcctaagtgtcctgggacaga
ccaggaggcacaagccactgtcccagccctctatctttttaattaaaaataaaacactataactcattgggaaacctg
ctgggacatttctataaaaaatgataaataagaaatgagtaattccacttttttttttttttttttttttttttttt
tcacccagggctggaatgcagtggtgcatctcagctctctcaaacctctgcctcctgggctcaggcactcctccacactc
agcctcccagtagccaggactacaggccactgccaccacacccagctaattttgtatttttttttttttttttttttt
acaggctgtgcccacacagccggctaattttgtatttttagtagagacgggttttctccatgttggtcacgctagtctcga
actcccaactccaggtaaactgcctgctcagcctcccaagtgctgggattacaggatgagccacgctgctggccaa
tttttttgtattttttaatagagacgggtttccactgttggccaggtggtctcgaaactcctgcaactcctgagtgaccgac
ccactcagactcccaagtgctgggattacaagcttgagccaccatgccagccatggcatgtgtttaatttttaaaaga
aactggcaaatgttttccaaaaatgactgtaacttacttaactccattagcagagataaagatttctgtgtgtctcaca
tcctcatcaacataaaaaatttgcaattttgtaaaaaacgctacacccatgctaaocaaaacaaagcaggggcaacggactacc
agtcttcagcctctcctctgagccatgccttgtagaggacaatttccacaacagctatcaagattcaaaatgcacacaca
aaaaaagagaaaaacacacacacacacacacacacacacacacacacagccaaactgggtgaaaccccatctctactaaaaata
caaaaaattagctgggcgtgtgtggcagggtgctgtaatgccagctacccaggagggttagaggcaggagaatcgttgaacc
cgggagggcagaggatgcagtgaggcgaagatagttgcaattgcaactccagcctggggccacagcgcgagactcgtactcaaaa
aagaaaaaagagaggagaaaaaaacacacaaagattacagatgcatatattctttgatttaggaaacccagttctggga
atttatcatttagataacttccccatgataactaatcatgagtgatagggattatttattgtgcaattctgtgtgtagc
aaaagattggaaaataactcaaatgtccataaataagaagacttcttaataaaatttggtacagacaataataattgttatag
gctttaaaaaagaaaaaaagagggaagtttcaatgtcctcccaagtagccaggattacaggcgtgagccacactgccc
gggataactttttgtatttttagtagagatgggttttggccatgtatacaagctgggtctccaggtccaaagcacaagcga
attccacccactcagcctcccaaaatgctaggattacaggcatgagctacogtgcocagcctgctcaaggaaagattttgt
atagctaggattctccagagaaaacagaactaggaagagggcagataggaagaactgacattataaaggaaattgactc
acacaaattatggagctgcagcagtcacaaagctgcaggttgggttggcaagttggagacccaggagagccaatgggtta
gttcaggtctgagttcccaatgctctgagagccagggaagtgaaatgataatctccaggtccaaagggcagcaggttgagac
caagggaacaaatcaatgtttcagtttgatcccaaggcgaagaaagactgagttccaggtccaaagcacaagcgaata
aaaaattctctcttatttggggaggggcagccttttctctattttaggtctcaactgactggatgtggcccaactcaatt
agagaggggaaatctgtttcactcactgctgattcttcttttcttttctttgttttggaaatttgaacttctcactctgtg
ccaggctggagtgccgtggcatgatctgggttccttgcaactccactccggggtcgaagcattctcctgctcagc
ttcctgagctagctgggttcaaggcaatcacacacaccccggttaattttgtatttttagtagagacgggtttcgcc
atgtgacacaggtgggtctgaatgctcaggtgattagcccgctcggtcccaaaagtgctgggattacaggc
gtgagccaacgtgcccggctcactgctgatttaaattgttaaccttagccgggtgcgggtgtttaccctgttttccca
gcactttgggagccagggggcggtcctcactgcagctcaggtgtgggtcaacaatgggtgaaactgctcacttaaat
acataaaaaattagctgggcatggtggcatgagcctgtagtccagctactcaggaggccagaggaacagagaatcactgaa

FIG. 7

55/157

ccacgagaggtggaggggtgcagtttaactgaaattgtgccactacactccagccaggggtgacaaaacaagactcttctctca
aaaaaagaagtagaaccaaatgtgcgtgttgttttatagtacatcaaaagctatccggaaggatcaatcaaaaactaaagat
agaggttctccatttagagatgtgcgtgggaataagaagactttcaacttttatgtgtttttttaaatacatatatattgtt
ttatctctcaaaaacttaagtaacaagaatttaaatacccttgcatgctctctctctgtagggtgtgaacctggag
ggaaacaaatcatgttttatactcaacttttgtattcctggattttagtacatagtaggtactcaataatagagggaagaa
cgaatgaatgaaccagaagatagcagatttagtgcgtgttagagagaggagggggccaggccagtggtccacacatgt
aatcccgagctatttgggaggtcgaggcaagaggtcactgaaacctaggagtggagacagcctgggtgacaaagcgaa
atctctgtctctccaaaaaataaaaaaataaaaaaagccaggagagagtggtccacactgttaattccagcga
ctttgggaggccgaggtgggtggatcacctgaggttagcagttcgagagcaactaggccaactgggtgaacacacatct
ctactgaagatcacaaaaatttagttggggcgtagggtgcgtgctgtaattccagcaactcaggaggtcggagcgaggaa
tagctgaaacctggggagaggttgcagtgagccagactcatgcccgtgcactccagcctggatgacagagttagaca
ctctataaaaaaataaaattcctggatattgggtggtatgcaacctgctcctcccaactattctagaggctcagggttgaaga
ttgctgcaacttgggaggtcgggtgtgcagtgagccaagattgtgccactgactccagcctggggcgacaggtaaaaac
ctgtctctaaaaaagaaggaggattagggtcgggtgcgtgaggtgtaattccagcaactttaggagggccggaggtc
gtgtggatcacctgaggtcaggagtttgagaccgctgacccaattagggtgaaccccgctctctactaaaaatacaaaaag
tgcgcgggtgtggtggcggtgtgctgttaattccagcacttgggagggtgagacagagggaattcactgaaaccagggaagt
ggagagtgcatgtgagcgagatcatacactgcaactcagcctgggtgacagagcaaatctctgtctcaaaaacaaaaa
aagaagaagaagaaattttttaaatttaataaaaaaagggaattataatcttatctaatgtaatcaaaaagttg
acatcctatcacatttgcgttatcttctacttatcagaagcaagtcactaagtcagccctcaactcaaggaaaaggttaca
tggggggtgacgtccaggaggaggaggttactgggagccaacttaggagctgtcactacaaggagggaagaaagcaatgaa
ggatgaaattctaggatgggtgttatctctgagtggaagcaagggggagagatgcaagaagcactggccgggtgaaagtta
tttgttaataaattccattcttggggcaggtgggtgatttatagggtgtcaaaagtagttgattggcaaaaatggggga
gagtgctatgaaccaggagttattttagtcgaattcagtaattttaaaatttcaataataatgaaataaaaaattgc
caactctcctgcaacttgcgtcttgagcccatcaatgcaactcagcttggtctagttaaacactggctccactgacctc
cacttgccccaaactaaccttagccctcacacacactgaccttgcctatgacagacagatctctccatagttgcaatt
aggcccccgaggttctattcctctgaaacttttatccatgtaacctcacctgtaggctgctatatgtaagaataggaa
tggtctgatggagtggagtcggcaggagacaggagacaggggcaggtcaacgttctgctcttagcactagcacttgggcccgaa
gacaggtttgatggctggttatacagcaactcactattggcctctacattggcaaggctggggcaggtctgtgagttct
tctctctttgaggtctctgctcgttatgatacagatattagctaaactgtgtcttgagaagtgcaactctgcacagcaac
gootgtaactccttctgttgataagatctcaactctctccttctctgttttgtagaagtagactaccatcactgga
gatggctatgtgtcaaggaaatgagagtggtctattggccaacaaatgagtaaaacactgaaagcttcaagtttgacagccca
ctaggaagtaaatcagtcagcaaccaagtaagcctggaagcctggatcctctcagttgtgctctcagatgagactgc
agttctattggcccttctgtgcagccttgtagggaaactgaaagtggaggaccaggaagcctggcctggaacttga
ccccagaaaattgtgagatataataatagttattgtttctaaagctgctaatgttggtataatttgttacaagtaag
agataatgtatacaggttctttctattcttttcttatgtaacagagttttctgtgcaggaggttggttggtctcagcc
tgtaatcccagcactttggaggccaaaggcaggtggatcaacttgaggtcaggagttcgagggcagcctagccaacatgggt
gaaacctgtctctcaaaaataaaaaataagcagatgtggtggcactgacctgtaattcacaatttgagaggtt
caggcaggagaaatcgtgtaacctcaggaggtggaggttgcaatggccagagatgctgcaattgcactccagcctggaaga
cagagcaagactccattttaaaaaaaacaaaacaaaagaagagttgtgctatgaaagttgtgtatacattgtctccca
gaacagaacaacacagagcctagggttctcctgggtatgtgtgggttatgctgaagattggaattgttagatctcagg
atgtttgttttgttttgagacagagtgctctctgttgctaggctggagtgcaagtggtggcatcttggtcactgcaac
ctctgcccccagcttcaagcaatctctcctgcagctccccagtagtggtggaactacaggtgccggccacacgccttg
gctaatgttgcgtattttataagagatgggttttaacattggccagcctgggttgtaactctggcctcaagtgaacc
acctgcttggtcctcccaagttgtgggattcacaggcctgagccaactgcccagctggatctcaggatttacaattgt
ccacattactggtagtcagaataacattaaagtttaggttcagcctgctgtagtgtgtggccagggttgggagttaa
gcttggtcctcaatccatggaagtgggattgtctgtgggcagggttgagtgctgtaacagagggccaggttaggggtcagtc
cttactgttaggttgagggtgagtggtggccacatttgggaacaaattgaagtctggtttggggttggttttggaacag
ttggagaacactgtggggtgggtttgggggtcaatcaaggccagggtgaagaattaggcttgctcatcctgtgcaaaac
tgacccctgggtgatcatcaactaccagctgtggccgcacagggaccagccaacaaagatgagatccactctggg
tcgaagaagctcttctcaatcaaggcgggggtggggagatggaggaagaagaaagaaactggctccctcagatgca
aaagagaaggcaggccagggaatagaacaacacactgacttaattggggcagccctgagccgtaacctctcaaacctcc
gctcagtagtaacccgttgcgcaggaggagggggagctgctacggggacatcgtcgatgtctcctcctactcctggcgatcc
cagcgtccacccctctctccagcgtctcggttctcgtcgccagagtggttccagggttaagctcccaggaggggcagg

FIG. 7

56/157

[illegible]

FIG. 7

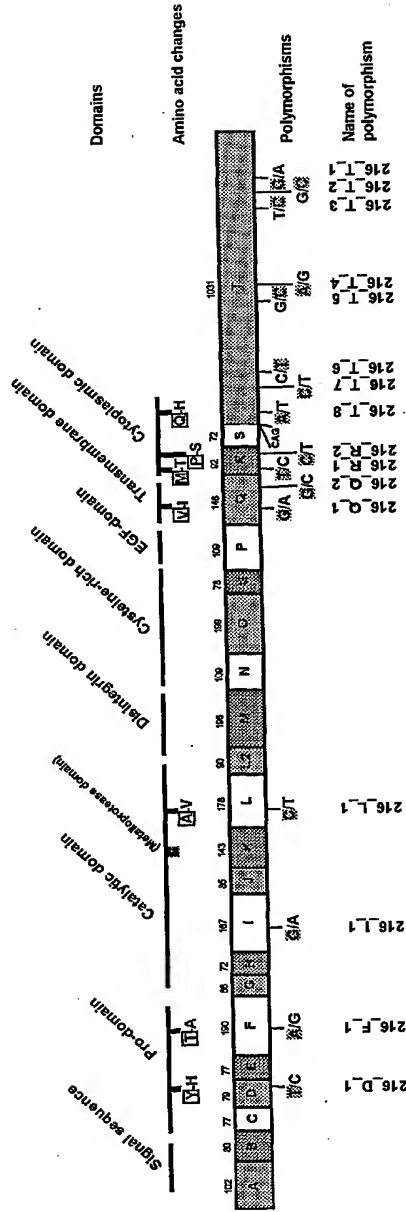


FIG. 8

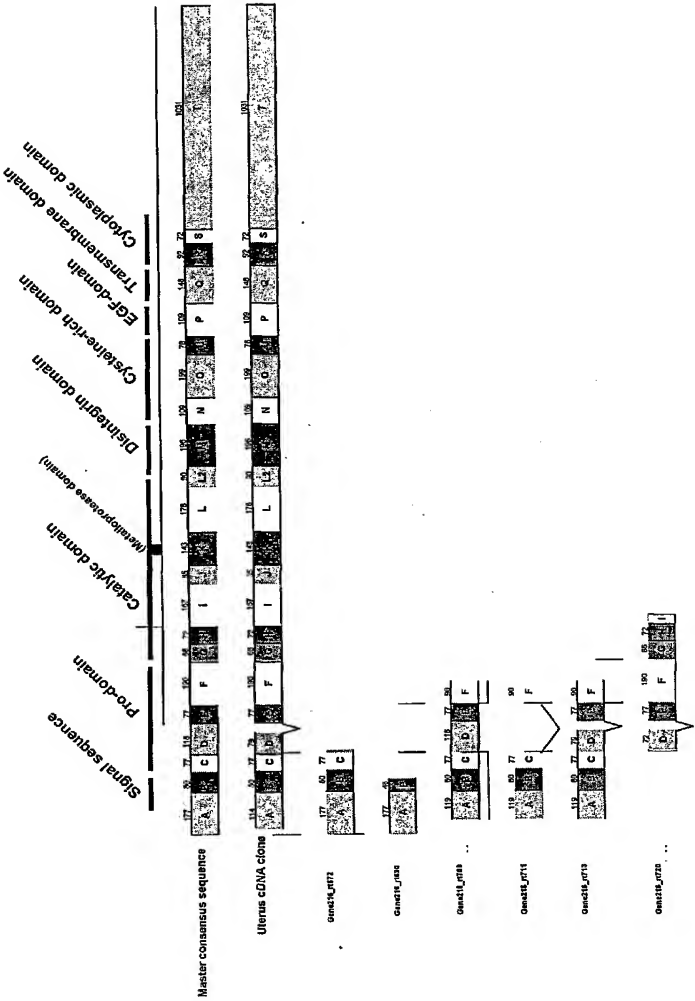


FIG. 9

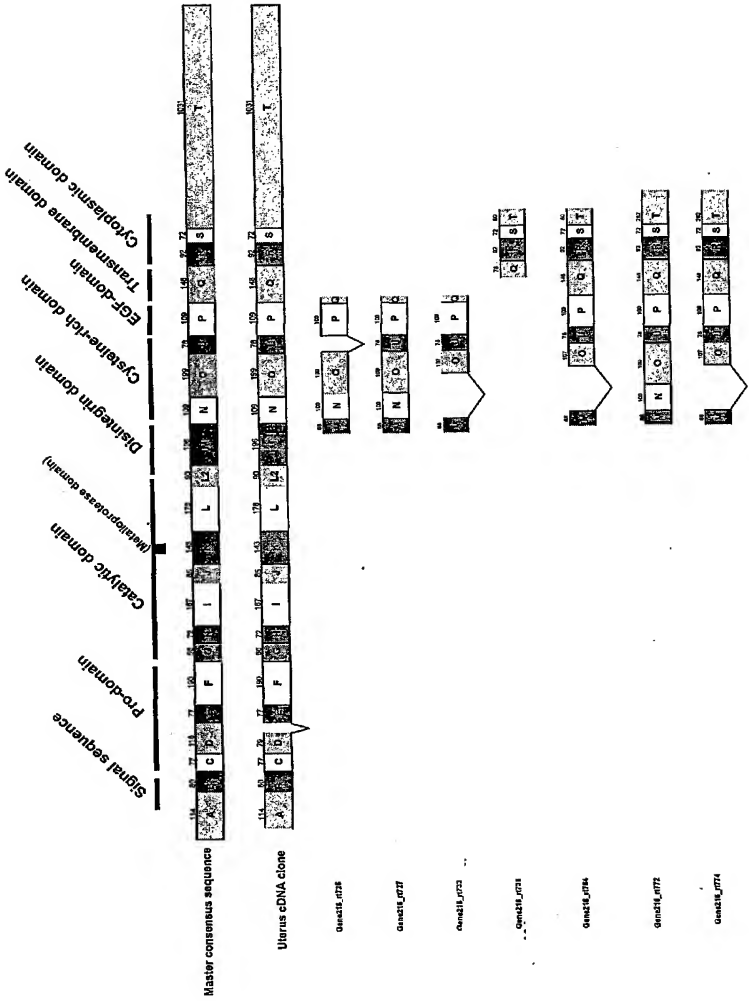
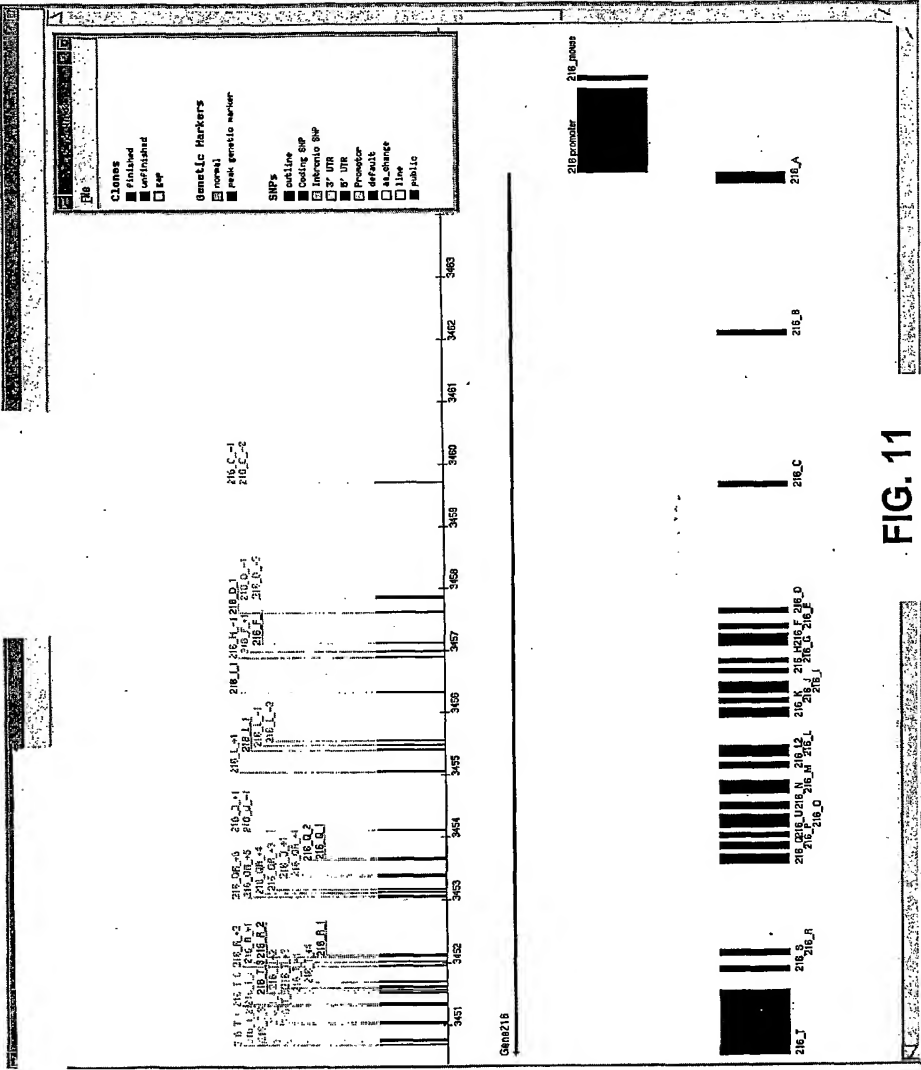


FIG. 10



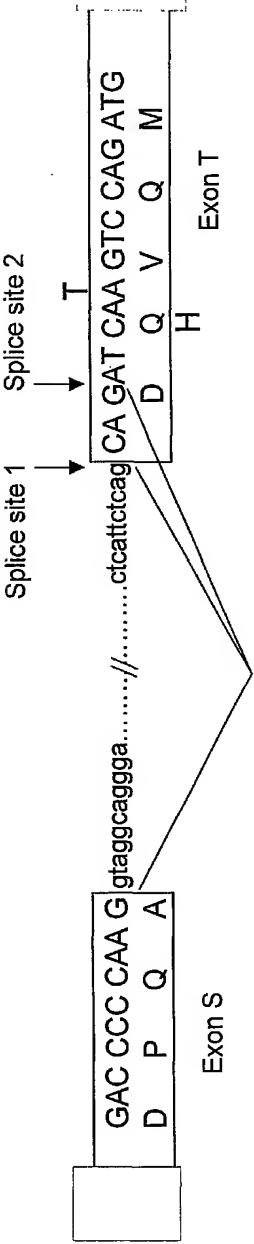


FIG. 12

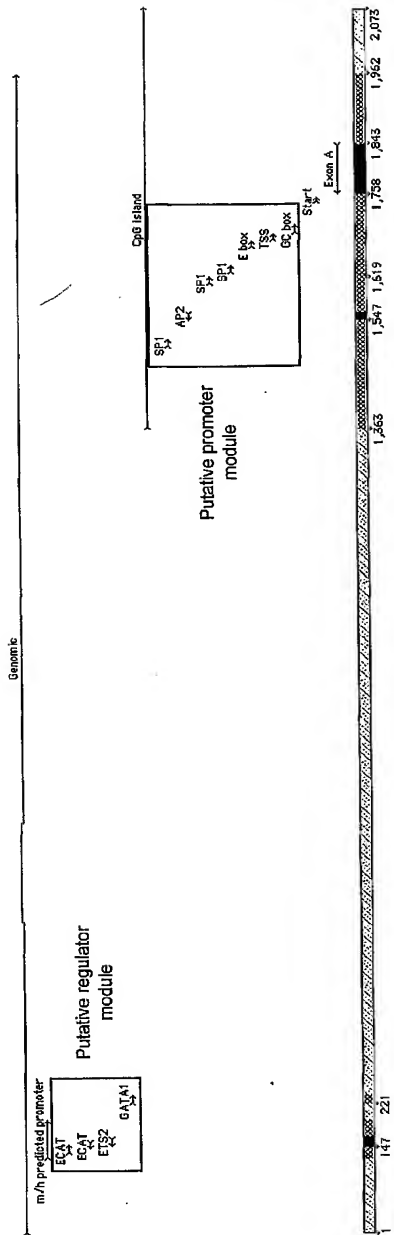


FIG. 13

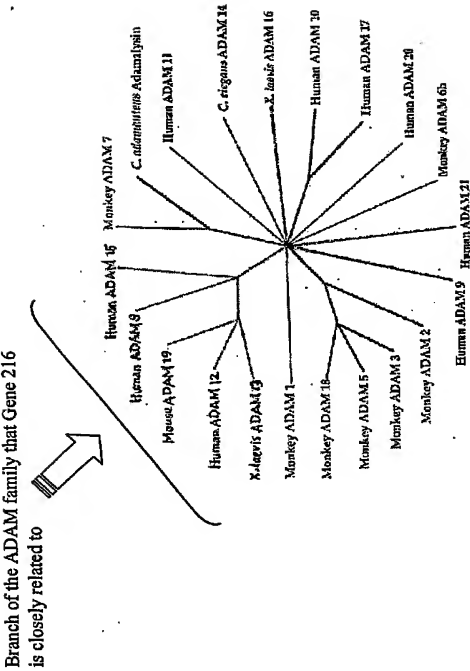


FIG. 14

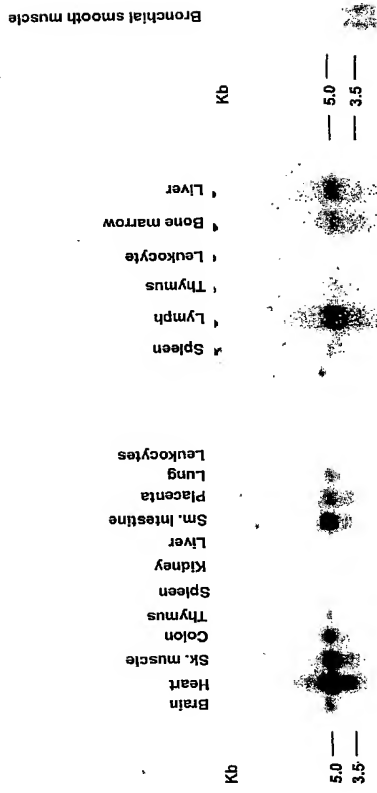


FIG. 15

	1	2	3	4	5	6	7	8
A	Whole brain	Amygdala	Caudate nucleus	Cerebellum	Cerebral cortex	Frontal lobe	Hippocampus	Medulla oblongata
B	Occipital lobe	Putamen	Substantia nigra	Temporal lobe	Thalamus	Nucleus accumbens	Spinal cord	
C	Heart	Aorta	Skeletal muscle	Colon	Bladder	Uterus	Prostate	Stomach
D	Testis	Ovary	Pancreas	Pituitary gland	Adrenal gland	Thyroid gland	Salivary gland	Mammary gland
E	Kidney	Liver	Small intestine	Spleen	Thymus	Peripheral leukocyte	Lymph node	Bone marrow
F	Appendix	Lung	Trachea	Placenta				
G	Fetal brain	Fetal heart	Fetal kidney	Fetal liver	Fetal spleen	Fetal thymus	Fetal lung	
H	Yeast total RNA 100ng	Yeast tRNA 100ng	E.coli rRNA 100ng	E.coli DNA 100ng	Poly r(A) 100ng	Human Cot1 DNA 100ng	Human DNA 100ng	Human DNA 500ng

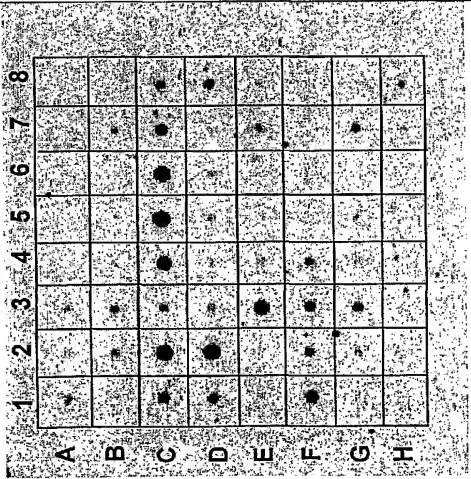


FIG. 16

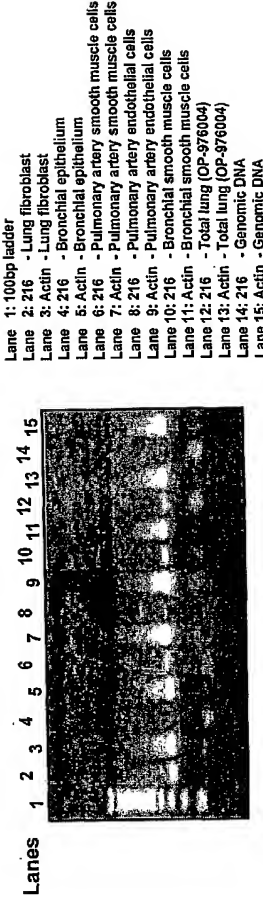


FIG. 17

Figure 18 displays the multiple sequence alignment of the protein sequences for adam12, adam19, gene216, adam15, adam9, and adam8. The alignment is organized into several domains: Signal sequence, Pro-domain, Catalytic Domain, Metalloprotease, Disintegrin Domain, Cysteine-rich Domain, EGF Domain, Transmembrane Domain, and Cytoplasmic Domain. The sequences are shown in black text on a white background, with gaps indicated by dashes. The alignment is presented in a grid format, with the domain names and residue numbers (00, 120, 160, 200, 240, 280, 320, 360, 400, 440, 480, 520, 560, 600, 640, 680, 720, 760, 800, 840, 880, 920, 960, 980, 1000) indicated above the sequences. The sequences are aligned in a way that shows the conservation of residues across the different proteins. The alignment is presented in a grid format, with the domain names and residue numbers (00, 120, 160, 200, 240, 280, 320, 360, 400, 440, 480, 520, 560, 600, 640, 680, 720, 760, 800, 840, 880, 920, 960, 980, 1000) indicated above the sequences. The sequences are aligned in a way that shows the conservation of residues across the different proteins.

FIG. 18

68/157

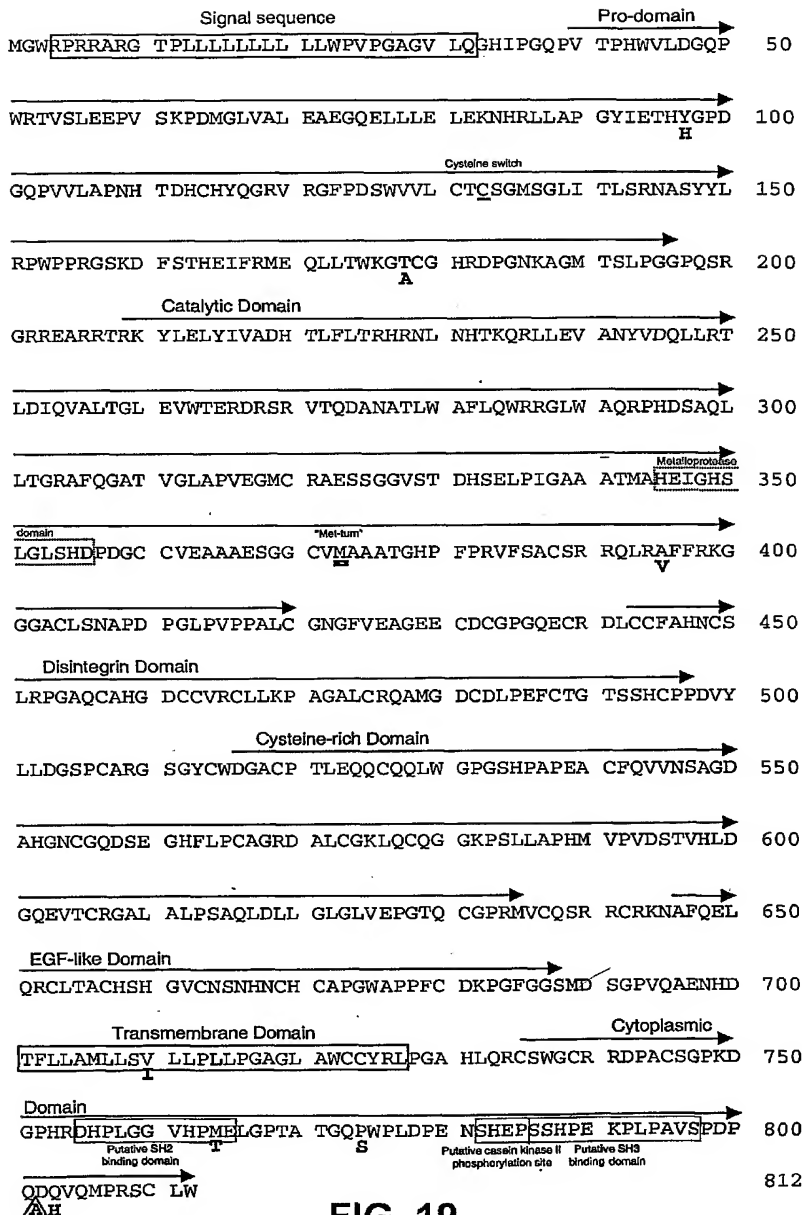


FIG. 19

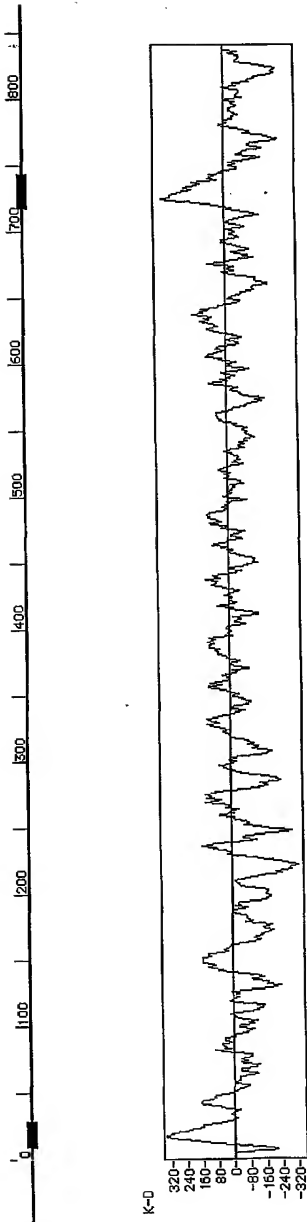


FIG. 20

70/157

ggagaatagtggggttgagtaggaggtagaaggttggaatgaattggaaaaggagtgag
ctgtgtagccaaggctgtgacctgtgctcattctaggccaatttcttcccttctgtctctcc
ctccctccctccctccctccctccctccctccctccctccctccctccctccctccctccctcc
ctcctagctatccttcccttcccttcccttacttctctcttcccttcttcttccctccctcc
tcccttctcttccccattccccctagaccaggtcccatgtagcccaaactgacttcaatc
tcattacagatgaccttgaacttcagatcctcagtcgccatttcttttctgtcctgaata
tcgtatttctctgtggagcctgtccagtatttctccaataaaatttcatccaggatac
aactggagtcgaatttcttcttctgcaacaaaaatgtctaggctagaatgtggagaattcaa
atattaaacgctagtgtctcagacaccatctgtgggcatgtggccagggagcataaatcg
cacagtgccttgcctcactgtctgtggaagtcacccggcccttccctgtcctcctatag
ctcctcacatgataactcacgggggtgggctgtcgtcagcagggaactggcacaggcttgg
cacatggcagggtggcacatggttcttgtacatggttgaatagaagcacagaaacccagct
ctaccagaggcaagacagggagagaaacaaacaaagatggcaagatgtccaggatgacc
gagagcaaaaggccacacctaagggtgggtggctgtgagggtttgggacatgtgacttg
aatccttctgtctcaagtacttccaagaggatgttggagacttctggggtggggtcctat
tgggcccattctcagtagtctaagccctcagaaaataattacttagctgtaattcaag
cacaatatatacacatggttaagtgtattagcaaaatggcttctctatgtaataaaataag
ataaaaaaacaaaacaaaacaaaacaaaacaaaacaaaacaaaacaaaacaaaacaaaac
tcccaaccctgcccagtgtagctgtacttactaatcaagttaatttaaaaatcaaaa
tataataccacagattttatttttccctgtgttctgtcttatacaaatggatcatattt
taaacatgagaacattgtgagatgactagccagtgctgtctgtctgtctctcctcagata
tttatttatttatttctgtatgttaggtacactgtcactgtctctcagacacaccagaaga
ggacattggatcccatcagatgggtgtgagccaccatgtgggtgtgagaaatgaact
caggacctctggaagtagttagtgccttaaccactgggcatcttccagtcctgac
cagtgctcttaattgtgaaatattcttcccttggccctagcatttctccttgggcatt
ttctcaggaaaaatgaaggcgctgtcacccctgagatttgttactaggatggtcataaca
gtgctacccatcagccaggaactgaaaacaaataaatgctcggcagcaggggaatgaccga
gtagacagagctacgcccacacagctatgagagctgtgcctacagctgagtggttctc
agacgggtcagtgggagggtgtgagcgatgggtgcctgtggcagttacatgtcctaccg
ctctattaaaggagctaaacaaaggcaaccttaaggaaaggaacatttatttgggtcacg
gttagtaggacagctctcagagtgggttgaaagatggcagcaagtgtctctagctgaggg
ggcaggagtgggaggccctggtcacttctgtgtcggcgagggaagcagacagagaggtgggg
gtgcccgttcttcttacttctccttctcctaccagtcaggatgacagaaacatcaggtat
ttgccaccacatttggagtggtcttctgtgtctgttaaaactctctggaaacacccct
cactgacctgtgtgcccgaacaagggtgattctcagcccaagcaagtcacagcgaagac
caaggacaggtatagtttaagggaatgggatagaagtcaatacacagccactggtgggcaaa
agggatgggaggagcacaatggagattcgagagccagggaagtgagcaggttactttg
tgaaaaatctcacatttctgaatctatatttggaaagaaaacttaatttaaaaaaatctg
ccacttctcctagcccttgtgtgtgtatcgacatgggagataccttgaggccaggaga
ctgagatcagcctgagcaacacaaataaacctcatttcttcaaaaacagaaacaaattggg
tgtgtgaggtgacgcctttaaaccagcaccagggaagcagaccaatcacccctgttaac
catctcaaatcagaccagggctgtctagcggcagaagatccatggatccactgtgac
catagtcatacaattggaactacttatcgacaagaataaatggccataggtaataggca
ataggttgggtaaaactcacaattgggaaaaatacaaccatactccagacaagggccaa
tcgcctttaaacatagaggtcctttaaacgtttaaaagaaggaccaaccactcagtaa
aaatgagctcacatgaatatgttactatatttaaaaaaaacccctgtgtctaaaa

FIG. 21

71/157

cagcagcttatgtaagaaattggtagtagtggctttctgtggggaaggagtagccaaagg
atggagactaaaggaagactctcctaagcattttataccttctcccttgtttgctttggt
gtagtgggtgggtgtttttgattctttttgtttgggtgggttttgggtttttcaaga
cagtgtttctttgtggagccctgggtgtcctggaactcactcaggctgccttgaactca
gagattctcctgctctgggcttaaaaggtgtacctcaccactgtcaggctgggttgggt
tttgagagcggggtcttgcctgtgtagctccagctgcctctgaagtgtcagaattacagaca
tgaactacagatatctcccagttctcctctcattctcctcctcctgcccccttgtatacac
attcacctgtgtgtgcatgtgtgactcactctgagttcttccagtggttactctccatcag
attttttgagatagggtctcttactgaacctggacctcaaacatttgctgtctctgtcc
agtccaaacacttggcttacaagtttgtgcaatgggttctggctttctctgattgctggaga
cttgaactcaagaaccaacgcagggtcctttacttactaagccttctccttgatgacaccc
ctcaccaccaactgccacaccagggtctccaggctgatctcacactttttccatataggac
tttgaactgtcttttgcctctctctgtgggtatcacaggtgtgtgctgtacacccaattta
tgggatgtctgggtaggaactcaggggcccaatcttgtatcaactgaggtgcgctgcctc
tctgcctctcattccc
cctccttcccttccctctctcttccatgagacttcttctatattgcacaggctagcta
tcttgagtgctcctgtctcgatccaattgcctgacctatcatactgggttttctgttttgg
ttttgggtttttcgagactgggtttctctgtgtagccctgggtgtcctggaactcacttt
gtagacagggtggcctcgaactcagaaatccgcctgcctctgcctcccgagtgctggga
ttaaaggcgtgcaccaccagccttttaatcccagcatttgatgttcgaatcgtataaatg
ttttttctctgcctgtggaacttttacattccatattctcagaacgaaattattaatca
tcttatttcaaagttaaatctcttgctatccgaaatcataccaagtgtacgaagtgatct
ttcaacacgaaagcaaatgtccttaaaatttgggaaagggtcctcatgctctgggggtct
gtggtatagaacaattctgtcctggccatcgctgttctctgcctgtgcccaccactgc
attgcctttaaccaggaaccacagacttcagcaacgcggccaccacaagaccctaggtac
ccagatgcatctcagctgcagtgaggacagcttcaactgctgtgggacactagtgacagta
gacagatcctgggtagaaaggagaatccagatgtcactagggagactgcgaccagagatg
tgtgctaggcagacagagcagctgtggggatcacttagctagtgcctctattcagcacca
gagaggttctetaagaaccttcccggactttccagcccatcagccatcccttggccagcg
tggataggattgttgcagataacgcacgggagggtcatcctggctcccgacgcgcacat
cctgttttgcgggcccagggttagggccacactcctgcgctcctgtgaggtggcgagggcc
cgctgagggcgggaggaggcgccccctctagccggggtggagccaaacagcacagcaaga
gtgagagggaaggcactcccagccatgggctcgaggtgcgggagaccgggggtctcc
gggtgctgctattgctgcgctgttgcctgcctcgtgtccgctgcggagcgctcggtgtt
tccaggtgaggatacagggcgggggggggggggggggggcggtcttagggcaggggtgaag
tgggtgggtcatgggttccagtgcaactacagaacatagagtttcccgaccctcggctac
ctggccaggacaggatgtgggaagggtctgtgctcctggtaagggtggccctggagcaga
cttggccacacttacttagcaaaagcaaataggcaacactaccctttcctcagtgtgccc
agctctgtccggaagcttctgcaacttccaaactgccaggccagagtgctcctcgtcct
ggaaggacactggggtccttggagtcctgggtctgcccacaggcctagctgtctcattcggc
tcgtcctctctcctatgctgggtctgggcagtggtatgggtgtttatacagttgccttatg
tggcagttcttcaaacaggggtgtctgcctagccgtgcctggggctaccacttctctctc
ctactgtaacctctccttactccgggacgacagggtcttagcacttccaaagcctgtagg
gtgaatctccaaactccacatagcaagcactaggctggccagaaggcagaaggaccacg
gggaggaggcagtaagggtgagggctggatattgggcagctctggaaccaccccatatc
tgttgactgctgcttccgtaaggcttctgagggggtgactgacagctgtattgccttta
gcctgcgcatagtttaattggtgaggagagagacctctgtgtttcctaattgcagactggct

FIG. 21

72/157

tcttcagtttccacttttagtccaccggactcctagcttgectccttcttctgtcttgagg
gctgcacagagaacactcacacctgaatccttatcctgcacccttggtaatccgaggcca
gcaagacttggttctccatactcagttaagccttctcttgccggagagatctttaacctc
tcaggaaggcttctctgtccttctctcctcctcctccttctctctcttcagcctccat
tccaaagccacacttgcacatccccacttgcacatggctcactaattacgaagcacatgt
gacttttaaggacgaggttctctggagctgctgctatcttcagatttggaaaaagtatatgag
cgttttgctcgcacatgcatctgtgtaccgcggaggtaccaggaaaacctgggggcccga
agagagaggcggtatccccggagctggggttacagatagttatgaggcccatgtgggtg
ctgagaactgaacccggctcctcagggaagaacagcgagtgctctcaatccccgatccacct
ctccagctcccattccagaatcttgatgtctgggacttagggaggcggtggaacaaggag
ctgggatgacccagggaagggtgagccaacagagctctccactcaggcaaggccagagct
ggagtggtgggtttctatgttttttttcttttttttctgttaacctgtgccaaactgga
agatgcagcgccagaggataagggttgctgctatctctgcaggagcttggtcctg
caggataggctcctgtgagtggttctgggaccagataggagatggggaggttaggaacattc
ctgaagaacattccaaacacgtctgagtaagtcttggtccgattctcttctggagccc
cttggaacgccttgcccaggatcactgagcctctgaacctagctccttcaggttctgg
gccaggtccgtcagagtcccaacaccagcatcaacctaatgggagtcaggagctgtggg
acagacatttgggtgctgtcttactcagatggcaaggagcctccaaacacctccctca
ccataaaggcatctactaccagctctaaaaataccaaggctgcctccttctccccacc
cctctgcgtgggttctgggttcttctccgggacctctgacctcaaacacagctctctaaag
ctctggccggttcaaatctcccaatggggatggttctcctggtgcctcgggttctctttgc
cacctcctcgtgtgcttctccatttaggggatgccccatgaataccaaatcctggaagt
ctacacatggatccagcttcttcaggaccctccttccccagtgctcctgtccttcttaact
gcaattttaacactattctctgctttcaccccaggaaatgccatggagagctagtcact
ccccactggatcctggaggccagactctggctcaagggtcacccctggaggagccggtaagt
atccggttatcctgtccttatttaggggacctaggtctcccgatgctccttctccccaa
atgcagcgctgacagtttttctgggttctcagaaccaaatactgtaccgttaaggcaact
ccctgttggtgggtgggacaaagagacatctagatggcgtgaggctgcacacatggaag
tcattttgggtctcctccctcccccccccttctattcctggccctaacctccttctccat
gggggaacattcaggccatgcccactcaactaagcaaatgaggtcatagtttggggcc
taaagcctttcaggagagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
gtgtgtgtgtattctgttctgttttgttttgattgagacagggtcttgctatgtagctct
ggggggcttcaagtcatgagccatcatgcccagccacaggagatttttaaaagaaggacac
caaaatgtctctcttggccgtgtttatggggaacatgtgactccgtggacatgctgctgt
cctgtaacctcagcttccatggctggaagccacctgggacaaaggaagctgtgggccattc
tctcctgtccccagatgtattctcactctgggttgcttctccagagcctgtcagcct
cctcctaactctgctcattgtgttccccagaaagagccatgtgtgccttcaactgcttcg
ctgcccattggggaggccggcactgatccgcctggcttacagagaggcaagcgagcccccac
agagcttgtaaaagtgtctcatgtctgaatagttcattgggacctagaaacctgaggtccg
tgggacctaaaggccttcagatgattggctgcctccctcttctcctaactactcaggc
tggagagaggtcaaatctgtacctcctgacagaggccagccccactgtccaaacctagag
tgttgcaacagtgaggatgcattcccagaaacctaaacctccaagttagggccttcaact
ggggtccccagacttctaggacttctctcaaccaagcaaatggttctgtggacatgca
cgggttccccgtatggtggccctctccttgcctggctcgtgtgagagatacccagg
gggactctgctctgcgttaaggacgtgtgaggccaggccagggaagagcttggcagggg
gtgtgtatttcacaaacaggggtgtactgcaggatggggatggtgcacaaagggggagg
agacctggagcagaagttaaaggcagcaaggccgcagaggggaaggggaccttccactggg

FIG. 21

73/157

gctactgcctttctccagagcagacattttcccataaagcaagaaggcactccaaccata
taagctcatgttttggcccttggagcctgtggcctggggaaagtggaggccactcttctct
acacctgctgcagaagggccctgaaagattcttggccagggtcccagcccagtgcatctt
gggataaaaaagggaaagccatcgtgggtggggaaacatttaaaaaaaaaatcacagcagc
ctccccctggaatctcttgggctagttccagttcttggcttctagccaggctaagtggacta
gcctgagagagaccaagtgcagtgagagagagagaggtggctagagggccaaggccagccct
tctgacatctagctaagagaggtcaacacttttttagggagccaaagttgggtttcatgttt
acttcatgagttcagttatgaggccagctcagaggaaatttcagaggatcgggcagtttgt
cagaactgaagggtggaggaaaaagtttgggttctctcagggaatgagggaaaaggcagcaag
atggagagatgtgcaaccaggagctgtggatgcagctgggtgatttgggtgcttgctggc
atgtgcaaggccctgggtttgatctatagcaccacagcaaaactagcaagcaagcaaaaca
agcatacaaacaaaacagaagcaagggttgcaaggtagtacaagagatggccaaagctgtag
cccacccaatgacccacatcaccctcagtgccactgctcctgcttttctgccttttc
atggggctcctgactatggccatagcatatgtccatagcagacataacaacatagccag
gcaacgtagcaactgctatgtccgtatatgttggtggttatgcatctttaagggctccgttgct
acaaacacaagcactgagaacatctctaagttacaactcccacacatatccactccctgc
aatgcgagagccagctgttccatgagctttctcactaggcagctctacatccctaccggt
tccctgggcagccaggtggccttggggcctagtgtgctcactgcgttcccttctcggtcagat
cttgaagcctgactcggtgctgtggtggtcttagaggctgaaggccaggtctcctgcttga
actggagaagaagcagtgagtaccagcgggggggtgctgaagtcagacaaaagaccctc
tctggagaggatcagtgctttctgggagggggttgggggctgggtgggaagcagcagtg
gagtgacaggggagagtgccggtacttggggctacagtggtgactgaaaccaactgtagg
aagtacatacctggctcatctcataccctggagcagctggatgggcggtctccccctggag
aagagtgagcaccagaaacagacattgactgataccttaattcaagggaaattcttaggca
aagggaaacttccacagatggcagaagaaaaagctgcgaaggtcaaattccaatggtgacag
ttcactctggtctacctggaatgcctccatctgcgtggaagagacagacagacatctggtct
acctggaatgcctccatgctgcgtggaagggaacagagaggagagggctgggggtggggggc
ttcccagtggggcccgatcgaagagggcattcttagtgctaaagatcttggtggacctaaa
gagtc aaagaacgggaggtggtaaaaggagaaacacagtgctgaaggagtcgacaggggtaa
gcaagatggaggcgtggaactgcagactccaccaccaaccacaccactcgtctatgtctg
ctttcttaaggacattgcagaaggcactcattcattttattctctccatggaatcagaca
gactgggaacaaaactgcaaacaccacgcacaaaccacagttcctgacagtgccacatccct
tctactaaacggctcttgtacatgtgtgcacacattgaacgttagcatgtatgagtgcag
ccagcaaaagcagagagttgtaattgtgtgtaaggatggctggccgtacaacagcctgggt
ttttgtatggaggacaatgtgaatttggagtcaggattttctgtgaagggaaggaattgaa
gtcaggcatggaggttagaggcctgtaagtcgggctcctctgggcaggagaaacacagagtc
aagtccaggactgctgagttataaagaatgagttcaagcagccctaggcaacttagca
aaatccagttctaaaatataaaggagtgaggaaagaggggaagagaaaaagaggggagaagg
agaaggagggaaaaagaaaaagaaaggaaggaagaaagaaagagaaggaaggaaggaag
gaagggaagataaaagaaagagagagacagacagacagaaagaaagacagaaagacagaaa
gaaagacagaaagaaagaaagaaagggcagaaaagcagctatttatttatggttgcaatgt
tccctggtggggacaggggaagagcctgaaaatggaaactactgatagagttctagctcc
aggtaaagggacccttccccgggtggctgaacagaagcagctacaggcaggtggaagtgcg
gaagtacgcatcactgaggaacagatggagacatgagatctagcagctctgaggagaactgc
gagggcctggggctccgctcagcaaaaggcaggagtaccctggccctacttggcctctctga
ggtagtatctcctggcac
acctgcctgcagcagggtccacagcaccaagggtcccagcctgtccttctgggtgcttc

FIG. 21

74/157

ccacacttcccatccccttgactaaccacctcaggactcgctccagacctttgtcttct
ttttgcctgaggtttttctcttgaggcacagtgtctcctttctctctgacacactgaatg
cctccgctagcactcaggggaggagctggaagcaagcggcaggatgcagggtctaaatcc
agtgaactcctaaccctcccccttccagcaagcttctggccccaggatatacagaaaacc
actacaggccagatgggcatccggtagtgctgtcccccaaccacacgggtgagacacttcc
atgggatctgagaacatggggaggggtgcagcctcaatgcctacccctgtcctgttttgcat
tcctaaggaggttgtagtctgccttcaggccccacttcctccttgtccccacacttgag
ctcctactcagatgtgggtatgagcctcttctcctctaggaatcattgccaatatocaggg
cgtgtgaggggcttccgggaatcctgggtggttctcagcacctgctctgggatgaggttaa
ggcgttgggggagaggggctgggctgcggcagaggtgagcgcttccctgttcacacatcc
ttgctatccactgcactctagtggccttattgtgtcagcagcaaatgcagctattatct
gcaacctcggaactcctggggataccaaagacttcccaaccacagagatcttccggatgga
gcagtgtgttcacctggagaggggtccagagagacaagaactcccaatacaaagcaggaat
ggccaggtcttccctcatgtccccagagccgggttaaggggtacagactagagtgggactgg
tccttcaaagcagaggacacaggaatgcgggtcctggagcggtctcctgttttaattc
agaggtgaagaatcaagtgttgagcagaagtttaactatccctcccttagcccgggggg
atcacggtgtgtggccaaggattcgggtcaagcagcctgccccctaccatttgctcctccc
aggtgagggcagagggcgcgaggagtcccaggtacctggaactgtacatagtggctgacc
acacctgggtgaggagagacctaaagaggtctgctgggtgggggactgctcagcacctgt
atgaccagggaacctctctgccttctcagttcttgcttcagcatcagaacttgaaaccacac
gagacagcgctcctggaggttgcaaatcggtggaccaggttggatgattagggcgctg
aaggttgggtaggtaggcagttgaaaagctctgggtgagtgctgggtcctggcttcagcag
gtctctccctcagatttctcaggactctggatatacagttgggtgttgacggggtggaagt
gtggaccgagcaggatctcagtcgcatcactcaggacgcaaacgaaacgctctgggctt
cctcaggtggcgccgggggtgtgggccaggagaccacacgactccacacaactgctcac
gtgggtgccatttac
caccgccgttgccgggatgtccttatttccaaactccacccagtcacgcctgctccatcat
ccttaggggcccgcaccttccagggtaccacggtgggcctggcacctgtggagggcatatg
ccgcgcggagagctccggaggttgtagcacaagtaagctctgctggggacctaggaggaga
gaggtttgggtagggtttgtggtcctcagtgaaaccggctcctcaggaccactcggaac
tccccatcggcacagcagccacctggccccagagataggccacagcctgggcctccacc
atgatccccagggtgctgctgagcaggccgatgcagagcaaggaggctgcgtcatggagg
cagccacagggtagcgaagaagggtcgttgttaccggacctagacaaagcgactgcgtat
ctttggttaccctcagttccctctttagtaagataaacagcctgctaataaggcctcgg
tgggaaagatgctgataaggtggagtctcagaatgactgcgggactagggtactccactc
tgactcttggtcctccttggggccggactgcocgattttagcgtcacagcctcctgtct
ccagttcctaggctcctctctctctcgattgctcgcttagcttagaaagaacagcacat
cttgagggtgtgtggctaaacctcaggggccagagtgagaaagttagcctgagcttaggag
gtggctccagcaccgagatgtgtccacaccaccaatgacatccatcttccatctgcag
gcaaccttcccgcgcttctcagcgctgcaagcgcgcagctgcgcaccttcttccg
caaagggggcggttcttgcctctccaaacacctcggcgccggggctcctggtgctgccag
ccgctgcggaaacggcttcttgaagcaggagaagagtgcgactgcgggtcttggccagggt
cagatcgtcatcttctcgtccctgggggtcaaggctaaccactgacatcttagcttagcagg
tctagcttgcctccaccatcacgtgttctgttctctcctttctgacttcagaagtgcacag
acccctgctgcttggccacaattgctccctgcgtgcgggggctcaatgtgccacgggtg
attgctgtgcgaggtgcctggtgaggatactggggagcgcctgacatcttagctgggc
tgggacttctagtcctttctgagtgtaggtttgaccacagtgatgggttcacgtagca

FIG. 21

75/157

ctccctcgggctttcagcaaagatctctctgctttcccttaatgggtctctgggtgtagtt
aaagtccgcgggcacgcctttgtctcctgctgagactgactgcgatactccccgagttctg
cacccggcacctccccgtattgccccgcagatgtttacctactggatggctcacccctgcgc
tgagggtcgcggctattgcctagacggctggtgtcccacgctggagcagcagtgccagca
gctatgggggctctggtgagaccgcacgctggtcctggtgcccgaaccaatactaaaacct
gcgggtttctactgagggcaagctccacccgtggaactgaggcccgagctgccgccttct
tactccccctccccccagggtccaagccggccccagagccatgtttccagcagatgaact
ccatggggaattcgcaagggaactgtggccaggaccacaagggtagcttccctgccttggtg
ctcagaggtgaggtgtgattgctgaggggtctgcagctgtaaagtaggggcgagcatgcgga
gggaacactccaagttgtgaccaccttccacttccctccccagggaagcctctggtggtgga
aactgctgtgccaggggaggggagccgaaccactagtgcgcacatagtactatggact
ccacaattctcctagaggccgcgaagtgggtttgcccaggggcctttgtgtctccagata
gtcacctggaccagcttgacttgggtctggttagagccaggcacccggctgtggacctagaa
tgggtgagccctgccaccaccaacccctcctggttatttagtctccatgccaaagtgttct
cctcactgccagtggggcacaatgcccataggtgtgccaggacaggcactgtcagaatgc
tacctcccaggagctggaaacgttgctgactgcctgccataacgggtgggtgagtagcct
aagggtcagggtgaaccttggaggtccttgctacctggtgacttttctatcctcatctta
ggtttgcataagcaatcgtaactgtcactgtgtgctggtggtcctccaccttctgtga
caagcctggccttgggtggtgagcgtggtgatagtggccctgcacagtctgcaagtatcccaat
gggtggggggcagcaggaaccacctgggcagtagcctgcttgagactcagcacccctgc
cctccacagaccagatgccttcccccttggccatgctcctcagcttccctgctcctctgc
tccccctgggcttgcctagcctgggtgctactaccagctcccaacattcttgcatacgaaag
gactgtgctgcaggaggaccaccttgggaataggtaggttccggtgctgagctctctctt
cctgagcctgccccatgggtcctgcttctcagaactctcagggtcttggtagagtga
ggctacagggagctggggctttaggaagctagatgggataccttattctctagatgtagt
agactccaggctgtgggaagaagtcogtgggtgtgtatcactgcctgaccagcacttg
gtggtgtggtcctttccatgggcttcccttgatttttttggttttatgttttttgggtgtg
ttgggttttttctgctttaaagaaattcaacacagcctgtgctccgctttgtcctgaaga
cgactctagcttcttctgtctcaatatcagccatctcccatggcctttgcccataattactc
ccttcaggctccctcggtgctagcccaggacccttgggtcctcctataaactggccagct
atagtgttgctcttctgtgcccaggactctgtgctctgtcccctgtagatcacagtgc
gtaaacccaattttctggcaggcagaaatgccacctgatagcacacattatctaccgagg
gcagcctcctcactagggccctcgggcaagtccaggcacattatgtccctgtctgaacttg
gaagtgttctagaccaattgaggggaaggcagggttgggttggagatgtgactagagggc
acctcaggccagaaacagcaccagcaggcccccaggagccagtgagacgggtctggggaaag
ccaggtagtgctggggggtggggcggtgtctgcagacagggtaacagggtggagtga
gttgggagggggcacttcagagggggtggcagctgcacacoggttatcgggatgggtgtca
aggacagtggggaatatctggatgaaccatccaaagaaatggcaaggctctgtgagaaaggt
ccccgcagttcagagctctagactggcagagcacggctaaacggagcatgggaaggcacag
ttcccaacaaaggagacatctctgcacttcagctcctaggtggggcccaacgggtgacgcctac
atctagactgagtgggggctggagtgagtcacctcgggagaaaggatgtccacagccct
ggaggccctgggacataagttaggtctggacatcttaaggacagaaacaggaaatgtggaat
tcgaagcttgagtggaatggagaagcaatcccccttggctcatcacacggctcactttca
accttctgaactcttattctggtctgtcactggccctgcagagacataccctggggcag
gtgcatccggtggagtttggctccatcatcactggagagccctcgccccctcgttaagtgt
gccccctgggaacatggagaggggaagcaagggtgggtgcttgcctatgcctccttctaa
atgctctccttaacacaccttcaatcctctgcctgctcagccccatggaccccttgcga

FIG. 21

76/157

acagcggttgcacccctccatctcttgacttgctctcagacccctgcgaactctgagcttac
ctaagaactaccctctgaagcagcctggctctacagattgagttccagacccctgccctatcc
ctatgggtatggaagcaccctgaggacccctcctgttgccagtcacctacctctgtctcagt
ttgttgtccctcctcagatttacagccttgcatcaataaagaaatgagacatgggcctc
agagaagctggtgcatagagaccatgatgctggaagccctaggggcagggaaggagac
actgtggttcttcttgggtccttatagagggaggacaaatgtgccctgccatgtgacttg
cagtcctcagtttctcagacgcactcttataatlcctatgggctgtatgctgagctctta
ctcagcataggaaccccagagcccgatcatgttgtatccgcctgccctgagagctgtgc
tattctgaaatggttagaatgtatctaataacaataaaatccacaagttatatcagtgttgt
tggctgtgacctgttaaaagggctctaagttgtttatataaaagatatggagatggattac
tgagaaagaaattagaaacaacatctggacagtggaggagccagcactggggaggaaagg
gcagacagatctctgagttcaaggccagcctggctacagattgagttccaggatagcta
ggccacacagagaagaaacccctgttttgaaccacaaacagtcataataaaaaacaaactt
agtagcaccttgtagacagacagagagaaagggtcccaggagagctcagacccagcaccag
aggtggcaaggcagagctcacaagcccagggtggcaaaagtgggatctgttagcataaacc
caaggggcgctgggaacaggcagggaagtaacccttgatgcaagctcactctggtgagggc
ccctcaccctggatgtctgttagcaagggaatcagtcagtgctcagtcctgttaacatc
tgtgagaggggaaaggctgctgcagacatggcctgagcagcatctggattcgaacatttg
cactttagggcctgctctctcccctgggtggggcactcgccattcatgtccttacgaca
gctgtgagagagaggttgctgcaagtggtatggctgggttagctcgagccaccaggga
atcatgatttccctcgacatctagtcattacaacacaggttcttttactacaatttta
actaccaatataaactaataatgtgtataaaatatataaacacagtacacatatataagta
catagattgtagtagacatatataatattatatatatataattgttaatttaattactt
aattttattttatcatttgatattttttactgttagttataacaatgtgcataatatat
gtgtaatatataaatataattttattatttaattattatatataaatttaatttaatttaa
ttatacctatatatttagtagacatatacataggttacagaatggctacaaaagtgccagga
gccatcaaggagagaagctaaaagccagcaagtgatcttctcgagcgggtctgccatggac
tgtacaattagtgatggatttgcttctgtaggcaaggacgaggagatttcattttaggaa
agattcctgctattaatatgcttttctcctggtattattaaatatataacaatcactagg
tattagcccaccgttttgaaacagaatgttctgcagaacaatgaagatgtactctcttgta
atgatgctatatagacaaatagattatttcttttttaaaaaagaaaaagagccggggcga
tgggtggcacatgcctttaatcccagcacttgaggaggcagaggcaggcagatttctgagtt
caaggccagcctggactacagagtgagttccaggacagcagggtcactcagagaaactc
tgtcttggaaaaaaaagagaggaagaaagaaaaagattttattttttttatatacat
atgagtacaccatcagacacacagaagagggcaccagacccattacagatggttggtga
gccaccatgtggttgctgggaattgaactcaggacctctggaagaacagttggtgctctt
aaccctgagccatctctccagcccaaatagatgatttcttaattcttaaggatgatcct
ataagaattcctaaacttacattagtaattattaagctcttttacaataggacttctatt
aagcttctctaatatgaaaaactcaataagaactctgcagctcttcaagtgtcatgagt
tagttgctcttgatagcaagtaggcatcaacaacttagagcacttctaggaggttggt
aaaaaccattaacccagtggtctttaaagggaactaacaataggctatagggtgcaaggaca
gaagataaaatattgactaggtttatcaatacaaaatttaccacaaaaggttatgtttt
gacttttcataaaaactctttatgaacctgtagaactggtgaaagatgacgaatgcttag
ccagataaattactoctaataagatatgcatgtgaatatctgtgctgtaaaactttttatg
tttgaacttccagtgaaactttgtttaaaaaaggggggggttgaaaaagccatgtgatc
tattctcctagaagggtacagaagactaagaaagattacattggagatgtaaccttgga
gagaaagctttgggagcaagagcatagagagcaaggccatttggcatcagagcaaggagg

FIG. 21

78/157

>mouse Gene 216 cDNA
 AAAGGCACTCCAGCCATGGGCTCGAGGTGCGGGAGACCCGGGGGTCTCCG
 GTGCTGCTATTGCTGCCGCTGTGCTGCCCTCGTGTCCGCTGCGGAGCGCTCG
 GATGTTTCCAGGAAATGCCCATGGAGAGCTAGTCACTCCCACTGGATCTCTGG
 AGGGCAGACTCTGGCTCAAGGTCACCCCTGGAGGAGCCGATCTTGAAGCCTGAC
 TCGGTGCTGCTGGCTTTAGAGGCTGAAGGCCAGGATCTCTGCTTGAAGCTGGA
 GAAGAAGCACAAAGCTTCTGCCCCAGGATACACAGAAACCACTACAGGCCAG
 ATGGGCATCCGTAGTGTCTGTCCCCCAACACACGGATCATTTGCCAATATCAC
 GGGCGTGTGAGGGGCTTCCGGGAATCCTGGGTGGTTCTCAGCACCTGCTCTG
 GGATGASTGCCCTTATTGTGCTCAGCAGCAAAGTCAGCTATTATCTGCAACCTC
 GGACTCCTGGGGATACCAAAGACTTCCCAACCCACGAGATCTTCCGGATGGAG
 CAGTTGTTCACTGGAGAGGGGTCCAGAGAGACAAGAACTCCCAATACAAAGC
 AGGAATGGCCAGTCTTCTCATGTCCCCAGAGCCGGGTGAGGCCAGAGGGCG
 CGCAGGAGTCCCAAGTACCTGGAACGTGTACATAGTGGCTGACACACCCCTGAA
 CTTGAACACACAGAGACAGCGCTCCTGGAGGTTGCCAATTGCGTGGACCAGA
 TTCTCAGGACTCTGGATATACACTTGGTGTGACCGGGCTGGAAGTGTGGACG
 GAGCAGGATCTCAGTCGCATCACTCAGGACGCAACGAAACGCTCTGGGCTTT
 CTTACAGTGGCGCCGCGGGGTGTGGGCCAGGAGACCACAGACTCCACACAA
 CTGCTCACGGGCCGCACCTTCCAGGGTACCACGGTGGGCCCTGGCACCTGTGG
 AGGGCATATGCCGCGCGGAGAGCTCCGGAGGTGTGAGCACAGACCACTCGGA
 ATCCCCATCGGCACAGCAGCCACCATGGCCCCAGAGATAGGCCACAGCCTGG
 GCCTCCACCATGATCCGAGGGCTGCTGCGTGCAGGCCGATGCAGAGCAAGG
 AGGCTCGCTCATGGAGGCAGCCACAGGGCACCTTTCCGCGCGCTCTTCAGCG
 CTTGCAGCGCGCCGCCAGCTGCGCACCTTCTTCCGCAAGGGGGCGGTCTTTCG
 CTCTCCAACACCTCGGCGCGCGGGGCTCCTGGTGTCTGCCAGCCGCTGCGGAA
 CGGCTTCTTGAAGCAGGAGAAAGAGTGGCACTCGGCTTCTGGCCAGAAAGTC
 CCGGACCCCTGCTGCTTTGCCCACAATTGCTCCCTGCGTGGGGGGCTCAATG
 TGCCCCAGGTGATTGCTGTGCGAAGTGCCCTGTAAAGTCCGCGGGCACGCTT
 TCGTCTCTGCTGCGACTGACTGCGATCTCCCGAGTTCTGCACCGGCACCTCC
 CCGTATTGCCCGCAGATGTTTACCTACTGGATGGCTCACCCTGCGCTGAGGG
 TCGCGGCTATTGCCCTAGACGGCTGGTGTCCCAGCTGGAGCAGCAGTGCCAGC
 AGCTATGGGGGCTGGTCCAGCGCGCCCCAGAGCCATGTTTCCAGCAGAT
 GAATCCATGGGGAATTCCGAAGGGAAGTGTGCCAGGACCACAAGGGTAGC
 TTCCTGCTTGTGCTCAGAGGGACGCTCTGTGTGGGAACTGCTGTGCCAGGG
 AGGGGAGCCGAACCACTAGTGCCGCACATAGTGACTATGGACTCCACAAATTC
 TCCTAGAGGGCGCGAAGTGGTTTCCGAGGGGCTTGTGCTCCAGATAGT
 CACCTGGACAGCTTGACTTGGGTCTGGTAGAGCCAGGCACCGGCTGTGGACC
 TAGAATGGTGTGCCAGGACAGGCACTGTGAGAAATGCTACCTCCAGGAGCTGG
 AACGTTGCTTGACTGCTGCCATAACGGTGGGCTTTCATAGCAATCGTAAGT
 GTCAGTGTGCTGCTGGCTGGGCTCCACCTTCTGTGACAAGCTGGCTTGGGT
 GGTAGGCTGGATAGTGCCCTGCACAGTCTGCAAAACGAGATGCTTCCCCCT
 GGCCATGCTCCTCAGCTTCTGTGCTGCTGCTGCTCCTGGGGCTGGCCTAGCCT
 GGTGTCTACTACAGCTCCCAACATCTGTGTATCGAAGGGGACTGTGTGCAGG
 AGGGACCCCTATGGAATAGAGACATACCCCTGGGCAGTGTGCATCCGTTGGA
 GTTTGGCTCCATCATCACTGGAGAGCCCTCGCCCCCTCCCCATGGACCTCTTG
 CCAACAGCGTTTCGCACCTCCATCTCTTGACTTGCTCTCAGACCTCGGAACCTC
 TGAGCTTACCTAAGAACTACCTCTGAAGCAGCTGCTACAGATTGAGTTCC
 AGACCTGCCCTATCCCTATGGTATGGAAGCACCTGAGGACCTCCTGTTGCCCA
 GTCACTTACCTCTGTCTCAGTTTGTGTCCCTCCTCAGATTTACAGGCTTGCA
 CAATTAAGAAATGAGACATGGGCTCAGAGAACTGTTGTATAGAGACCATG
 ATGCTGGAACCCCTAGGGGCAGGGAAGGAGACACTGTGTTCTTCTTGGGTC
 CTTATAGAGGAGGACAAATGTGCCCTGCCATGTGACTTGCAGTCTCAGTTTC
 TCAGACCACTCTTATAATTCTATGGGCTGTATGCTGAGCTCTTACTCAGCATA
 GGAACCCAGAGCCCGATCATGTGTTGATCCSCCTGCCCTGAGAGCTGTGCTAT
 TCTGAAATGTTAGAATGTATCTAATAACAAATAATCCACAGTTATATCAGHAA
 AAAAAAA

FIG. 22

79/157

>mouse Gene 216_protein
MGSRCGRPGGSPVLLLLPLLLPSCPLRSARMFPGNAHGELVTPHWILEGRWLKVTLEEPILKPDSDVLVALEAEGQDLLL
ELEKKHKLLAPGYTETHYRPDGHFVVLSPNHTDHCQYHGRVRFRESWVVLSTCSGMSGGLIVLSSKVSYYLQPRTPGDTK
DFPTHEIFRMEQLFTWRGVQRDKNSQYKAGMASLPHVPQSRVRREARRSPRYLELYIVADHTLNLNHTRQLLEVANCVD
QILRTLDIQLVLTGLEVWTEQDLSRITQDANETLWAFLOWRRGVWARRPHDSTQLLTGRTFQGTTVGLAPVEGICRAESS
GGVSTDHSELPIGTAAATMAHEIGHSLGLHHDPEGCCVQADAEQGGCVMEAATGHFFPRVFSACSRRLRTFFRKGGGPCL
SNTSAPGLLVLPSCNGGFLEAGEECDCCSGGQKCPDPCCFAHNC SLRAGAQAHGDCCAKCLLKSAGTFCRPAATDCCDLP
EFCTGTSPYCPADVLYLLDGSPCAEGRGYCLDGWCPTLEQQCCQLWGPGSKPAPEPCFQQMNSMGNSQGNCGQDHKGSFLP
CAQRDALCGKLLCQGGEPNPLVPHIVTMDSTILLEGREVVCRGAFVLPDSHLDQLDLGLVEFGTGCGPRMVCQDRHCQNA
TSQELERCLTACHNGVCN SNRNCHCAAGWAPPFCDKPGLGGSVDSGPAQSANRDAFFIAML LLSFLLPLLPAGLAWCY
QLFTFCHRRGLCCRRDPLWNRDIPLGSVHPVEFGSIITGEPSPFPFPWTS CQQRSHPPSLDLLSDPANSELT

FIG. 22

[illegible]

FIG. 23

81/157

```

mGene216 582 VPHIVTMDSTILLEGREVVCRGAFVLPDShLDQLDLGLVEPGTGCGPRMV 631
      ||.|. .|||: |:|. ||||| |||. || ||||| |||||
hGene216 600 APHMPVVDSTVHLDGQEVTCRGALALPSAQLDLLGLGLVEPGTQCGPRMV 649
      ||.|. .|||: |:|. ||||| |||. || ||||| |||||
mGene216 632 CQDRHCQNATSOELERCLTACHNGGVCNSNRNCHCAAGWAPPFCDKPGLG 681
      ||.|. .|||: |:|. ||||| |||. || ||||| |||||
hGene216 650 CQSRRCRKNAFQELQRCLTACHSHGVCNSNHNCHCAPGWAPPFCDKPGFG 699
      ||.|. .|||: |:|. ||||| |||. || ||||| |||||
mGene216 682 GSVDSGPAQSANRDAFFLAMLLSFLPLLLPGAGLAWCYQLP.TFCHRRG 730
      ||.|. .|||: |:|. ||||| |||. || ||||| |||||
hGene216 700 GSMDSGPVQAENHDTFLLAMLLSVLLPLLLPGAGLAWCCYRLPGAHLQRCS 749
      ||.|. .|||: |:|. ||||| |||. || ||||| |||||
mGene216 731 LCCRRDPLW.....NRDIPLGSVHPVEFGSIITGEPSPPPPWTSQQR 773
      ||||| .|| ||||| |||. || ||||| |||||
hGene216 750 WGCRRDPACSGPKDGPHRDHPLGGVHPMELGPTATGQPWPLDPENSHEPS 799
      ||||| .|| ||||| |||. || ||||| |||||
mGene216 774 SHP..PSLDLLSDPANSELT..... 791
      ||| |. || :.
hGene216 800 SHPEKPLPAVSPDPQADQVQMPSCLW 826

```

FIG. 23

82/157

```

      10              30              50
CGGGCACGGGTCGGCCGCAATCCAGCCTGGGCGGAGCCGGAGTTGCGAGCCGCTGCCTAG
-----+-----+-----+-----+-----+-----+-----+
      70              90              110
AGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCCCGTTGC
-----+-----+-----+-----+-----+-----+-----+
                        MetGlyTrpArgProArgArgAlaArgGlyThrProLeuL

      130              150              170
TGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGTGCTTCAAGGAC
-----+-----+-----+-----+-----+-----+-----+
euLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGlyH

      190              210              230
ATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCTGGATGGACAACCTGGCGCACCG
-----+-----+-----+-----+-----+-----+-----+
isIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpArgThrV

      250              270              290
TCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAAG
-----+-----+-----+-----+-----+-----+-----+
alSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluAlaGluG

      310              330              350
GCCAGGAGCTCCTGCTTGAGCTGGAGAAGAACCACAGGCTGCTGGCCCCAGGATACATAG
-----+-----+-----+-----+-----+-----+-----+
lyGlnGluLeuLeuLeuGluLeuGluLysAsnHisArgLeuLeuAlaProGlyTyrIleG

      370              390              410
AAACCCACTACGGCCCAGATGGGCAGCCAGTGGTGTGGCCCCCAACCACACGGTGAGAT
-----+-----+-----+-----+-----+-----+-----+
luThrHisTyrGlyProAspGlyGlnProValValLeuAlaProAsnHisThrValArgC

      430              450              470
GCTTCCATGGGCTCTGGGATGCACCGCCAGAGGATCATTGCCACTACCAAGGGCGAGTAA
-----+-----+-----+-----+-----+-----+-----+
ysPheHisGlyLeuTrpAspAlaProProGluAspHisCysHisTyrGlnGlyArgValA

```

FIG. 24

83/157

```

      490              510              530
GGGGCTTCCCCGACTCCTGGGTAGTCTCTGCACCTGCTCTGGGATGAGTGGCCTGATCA
-----+-----+-----+-----+-----+-----+
rgGlyPheProAspSerTrpValValLeuCysThrCysSerGlyMetSerGlyLeuIleT

      550              570              590
CCCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGGCCACCCGGGGCTCCAAGGACT
-----+-----+-----+-----+-----+-----+
hrLeuSerArgAsnAlaSerTyrTyrLeuArgProTrpProProArgGlySerLysAspP

      610              630              650
TCTCAACCCACGAGATCTTTCCGATGGAGCAGCTGCTCACCTGGAAGGAACCTGTGGCC
-----+-----+-----+-----+-----+-----+
heSerThrHisGluIlePheArgMetGluGlnLeuLeuThrTrpLysGlyThrCysGlyH

      670              690              710
ACAGGGATCCTGGGAACAAAGCGGGCATGACCAGCCTTCCTGGTGGTCCCCAGAGCAGGG
-----+-----+-----+-----+-----+-----+
isArgAspProGlyAsnLysAlaGlyMetThrSerLeuProGlyGlyProGlnSerArgG

      730              750              770
GCAGGCGAGAAAGCGCGCAGGACCCGGAAGTACCTGGAACGTACATTGTGGCAGACCACA
-----+-----+-----+-----+-----+-----+
lyArgArgGluAlaArgArgThrArgLysTyrLeuGluLeuTyrIleValAlaAspHist

      790              810              830
CCCTGTTCTTGACTCGGCACCGAAACTTGAACCACACCAACAGCGTCTCCTGGAAGTCG
-----+-----+-----+-----+-----+-----+
hrLeuPheLeuThrArgHisArgAsnLeuAsnHisThrLysGlnArgLeuLeuGluValA

      850              870              890
CCAACTACGTGGACCAGCTTCTCAGGACTCTGGACATTGAGTGGCGCTGACCGGCCCTGG
-----+-----+-----+-----+-----+-----+
laAsnTyrValAspGlnLeuLeuArgThrLeuAspIleGlnValAlaLeuThrGlyLeuG

      910              930              950
AGGTGTGGACCGAGCGGGACCGCAGCCGCGTCACGCAGGACGCCAACGCCACGCTCTGGG
-----+-----+-----+-----+-----+-----+
luValTrpThrGluArgAspArgSerArgValThrGlnAspAlaAsnAlaThrLeuTrpA

      970              990             1010

```

FIG. 24

84/157

CCTTCCTGCAGTGGCGCCGGGGCTGTGGGCGCAGCGCCCCACGACTCCGCGCAGCTGC
-----+-----+-----+-----+-----+
1aPheLeuGlnTrpArgArgGlyLeuTrpAlaGlnArgProHisAspSerAlaGlnLeuL

1030 1050 1070
TCACGGGCGCGCCTTCCAGGGCGCCACAGTGGGCCTGGCGCCCGTCGAGGGCATGTGCC
-----+-----+-----+-----+-----+
euThrGlyArgAlaPheGlnGlyAlaThrValGlyLeuAlaProValGluGlyMetCysA

1090 1110 1130
GCGCCGAGAGCTCGGGAGGCGTGAGCACGGACCACTCGGAGCTCCCCATCGGCGCCGCAG
-----+-----+-----+-----+-----+
rgAlaGluSerSerGlyGlyValSerThrAspHisSerGluLeuProIleGlyAlaAlaA

1150 1170 1190
CCACCATGGCCCATGAGATCGGCCACAGCCTCGGCCTCAGCCACGACCCCGACGGCTGCT
-----+-----+-----+-----+-----+
1aThrMetAlaHisGluIleGlyHisSerLeuGlyLeuSerHisAspProAspGlyCysC

1210 1230 1250
GCGTGGAGGCTGCGGCCGAGTCCGGAGGCTGCGTCATGGCTGCGGCCACCGGGCACCCGT
-----+-----+-----+-----+-----+
ysValGluAlaAlaAlaGluSerGlyGlyCysValMetAlaAlaAlaThrGlyHisProP

1270 1290 1310
TTCCGCGCGTGTTCAGCGCCTGCAGCCGCCGCCAGCTGCGCGCCTTCTTCCGCAAGGGGG
-----+-----+-----+-----+-----+
heProArgValPheSerAlaCysSerArgArgGlnLeuArgAlaPhePheArgLysGlyG

1330 1350 1370
GCGGCGCTTGCTCTCCAATGCCCGGACCCCGGACTCCCGGTGCCGCCCGCGCTCTGCG
-----+-----+-----+-----+-----+
lyGlyAlaCysLeuSerAsnAlaProAspProGlyLeuProValProProAlaLeuCysG

1390 1410 1430
GGAACGGCTTCGTGGAAGCGGGCGAGGAGTGTGACTGCGGCCCTGGCCAGGAGTGCCGCG
-----+-----+-----+-----+-----+
lyAsnGlyPheValGluAlaGlyGluGluCysAspCysGlyProGlyGlnGluCysArgA

1450 1470 1490

FIG. 24

85/157

```

ACCTCTGCTGCTTTTGCTCACAAGCTGCTCGCTGCGCCCGGGGGCCAGTGCGCCACGGGG
-----+-----+-----+-----+-----+-----+
spLeuCysCysPheAlaHisAsnCysSerLeuArgProGlyAlaGlnCysAlaHisGlyA

      1510              1530              1550
ACTGCTGCGTGCCTGCTGCTGAAGCCGGCTGGAGCGCTGTGCCGCCAGGCCATGGGTG
-----+-----+-----+-----+-----+
spCysCysValArgCysLeuLeuLysProAlaGlyAlaLeuCysArgGlnAlaMetGlyA

      1570              1590              1610
ACTGTGACCTCCCTGAGTTTTCACGGGCACCTCCTCCCACTGTCCCCAGACGTTTACC
-----+-----+-----+-----+-----+
spCysAspLeuProGluPheCysThrGlyThrSerSerHisCysProProAspValTyrL

      1630              1650              1670
TACTGGACGGCTCACCTGTGCCAGGGGCAGTGGCTACTGCTGGGATGGCGCATGTCCCA
-----+-----+-----+-----+-----+
euLeuAspGlySerProCysAlaArgGlySerGlyTyrCysTrpAspGlyAlaCysProT

      1690              1710              1730
CGCTGGAGCAGCAGTGCCAGCAGCTCTGGGGGCCTGGCTCCCACCCAGCTCCCGAGGCCT
-----+-----+-----+-----+-----+
hrLeuGluGlnGlnCysGlnGlnLeuTrpGlyProGlySerHisProAlaProGluAlaC

      1750              1770              1790
GTTTCCAGGTGGTGAACCTCTGCGGGAGATGCTCATGGAACTGCGGCCAGGACAGCGAGG
-----+-----+-----+-----+-----+
ysPheGlnValValAsnSerAlaGlyAspAlaHisGlyAsnCysGlyGlnAspSerGluG

      1810              1830              1850
GCCACTTCCTGCCCTGTGCAGGGAGGGATGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTG
-----+-----+-----+-----+-----+
lyHisPheLeuProCysAlaGlyArgAspAlaLeuCysGlyLysLeuGlnCysGlnGlyG

      1870              1890              1910
GAAAGCCCAGCCTGCTCGCACCGCACATGGTGCCAGTGGACTCTACCGTTCACCTAGATG
-----+-----+-----+-----+-----+
lyLysProSerLeuLeuAlaProHisMetValProValAspSerThrValHisLeuAspG

      1930              1950              1970
GCCAGGAAGTGACTTGTGCGGGAGCCTTGCGCACTCCCCAGTCCCAGCTGGACCTGCTTG

```

FIG. 24

86/157

```

-----+-----+-----+-----+-----+
lyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLeuAspLeuLeuG
      1990              2010              2030
GCCTGGGCGCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAATGGTGTGCCAGAGCAGGC
-----+-----+-----+-----+-----+
lyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCysGlnSerArgA

      2050              2070              2090
GCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCCTGACTGCCCTGCCACAGCCACG
-----+-----+-----+-----+-----+
rgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCysHisSerHisG

      2110              2130              2150
GGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCCACCCTTCTGTG
-----+-----+-----+-----+-----+
lyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaProProPheCysA

      2170              2190              2210
ACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGAAACCATGACA
-----+-----+-----+-----+-----+
spLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGluAsnHisAspT

      2230              2250              2270
CCTTCCTGCTGGCCATGCTCCTCAGCGTCCTGCTGCCTCTGCTCCCAGGGGCCGGCCTGG
-----+-----+-----+-----+-----+
hrPheLeuLeuAlaMetLeuLeuSerValLeuLeuProLeuLeuProGlyAlaGlyLeuA

      2290              2310              2330
CCTGGTGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTGGGGCTGCAGAA
-----+-----+-----+-----+-----+
laTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTrpGlyCysArgA

      2350              2370              2390
GGGACCTTGCGTGCAGTGGCCCCAAAGATGGCCACACAGGGACCACCCCTGGGCGGCG
-----+-----+-----+-----+-----+
rgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisProLeuGlyGlyV

      2410              2430              2450
TTCACCCCATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCTGGACCTGAGA
-----+-----+-----+-----+-----+

```

FIG. 24

87/157

alHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLeuAspProGluA

ACTCTCATGAGCCCGAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTCGCCTGACCCCC
-----+-----+-----+-----+-----
snSerHisGluProSerSerHisProGluLysProLeuProAlaValSerProAspPROG

 2530 2550 2570

AAGCAGATCAAGTCCAGATGCCAAGATCTGCGCTCTGGTGAGAGGTAGCTCCTAAAATGA
-----+-----+-----+-----+-----+-----+
lnAlaAspGlnValGlnMetProArgSerCysLeuTrpEnd

2590 2610 2630
ACAGATTTAAAGACAGGTGGCCACTGACAGCCACTCCAGGAACTTGAAGTGCAGGGGCAG

2650 2670 2690
AGCCAGTGAATCACCGACCTCCAGCACCTGCAGGCAGCTTGGAAAGTTTCTTCCCGAGT

2710 2730 2750
GGAGCTTCGACCCACCCACTCCAGGAACCCAGAGCCACAATTAGAAGTTCCTGAGGGCTGG

2770 2790 2810
AGAACACTGCTGGGCACACTCTCCAGCTCAATAAACCATCAGTCCCAGAAGCAAAGGTCA

2830 2850 2870
CACAGCCCCCTGACCTCCCTCACCAGTGGAGGCTGGGTAGTGCTGGCCATCCCCAAAGGCC

2890 2910 2930
TCTGTCTCTGGGAGTCTGGTGTGTCCTCTACATGCAATTCCACGGACCCAGCTCTGTGGA

2950 2970 2990
GGGCATGACTGCTGGCCAGAGCTAGTGGTCCTGGGGCCCTATGGTTCGACTGAGTCCAC

FIG. 24

88/157

```

3010          3030          3050
ACTCCCCCTGCAGCCTGGCTGGCCTCTGCAAACAAACATAATTTTGGGGACCTTCCTTCCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+
3070          3090          3110
GTTTCTTCCCACCCCTGTCTTCTCCCCCTAGGTGGTTCCTGAGCCCCCACCCTCAATCCCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+
3130          3150          3170
TGCTACACCTGAGGTTCTGGAGCTCAGAATCTGACAGCCTCTCCCCCATTTCTGTGTGTGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+
3190          3210          3230
CGGGGGGACAGAGGGGAACCATTTAAGAAAAGATACCAAAGTAGAAGTCAAAAGAAAGACA
-----+-----+-----+-----+-----+-----+-----+-----+-----+
3250          3270          3290
TGTTGGCTATAGGCGTGGTGGCTCATGCCCTATAATCCCAGCACTTTGGGAAGCCGGGGTA
-----+-----+-----+-----+-----+-----+-----+-----+-----+
3310          3330          3350
GGAGGATCACCAGAGGCCAGCAGGTCCACACCAGCCTGGGCAACACAGCAAGACACCGCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+
3370          3390          3410
TCTACAGAAAAATTTTAAATTAGCTGGGCGTGGTGGTGTGTACCTGTAGGCCTAGCTGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+
3430          3450          3470
TCAGGAGGCTGAAGCAGGAGGATCACTTGAGCCTGAGTTCAACACTGCAGTGAGCTATGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+
3490          3510          3530
TGGCACCCTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCCTCAAATAAATTTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+
3550          3570          3590

```

FIG. 24

89/157

AAAAAGACATAA
-----+-----+-----+-----+-----+-----+

3610

AAAAAAAAAAAAAAAAAAAAAAAAA
-----+-----+-----

FIG. 24

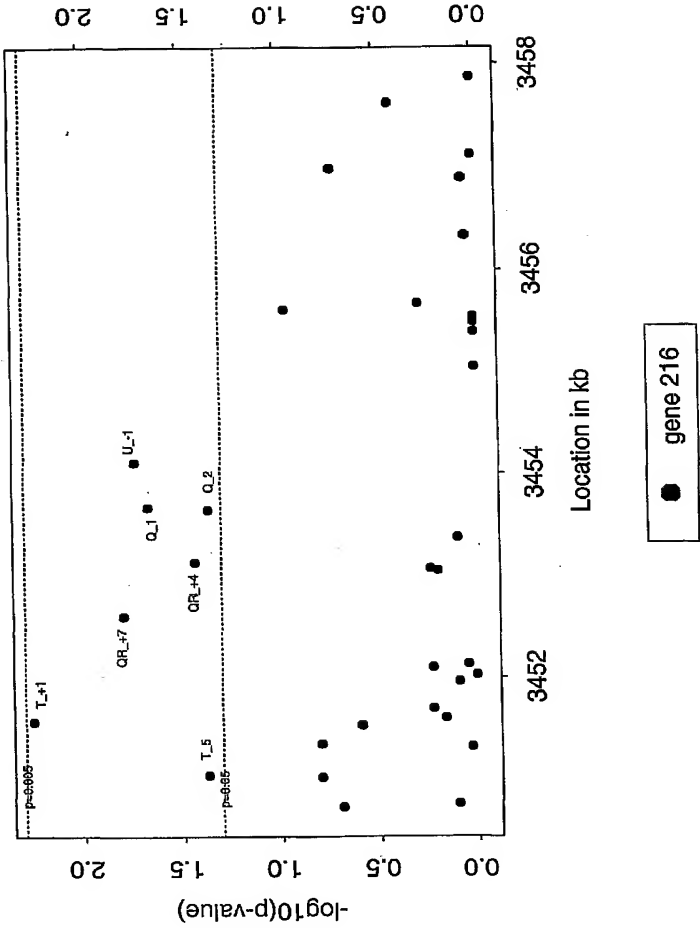


FIG. 25

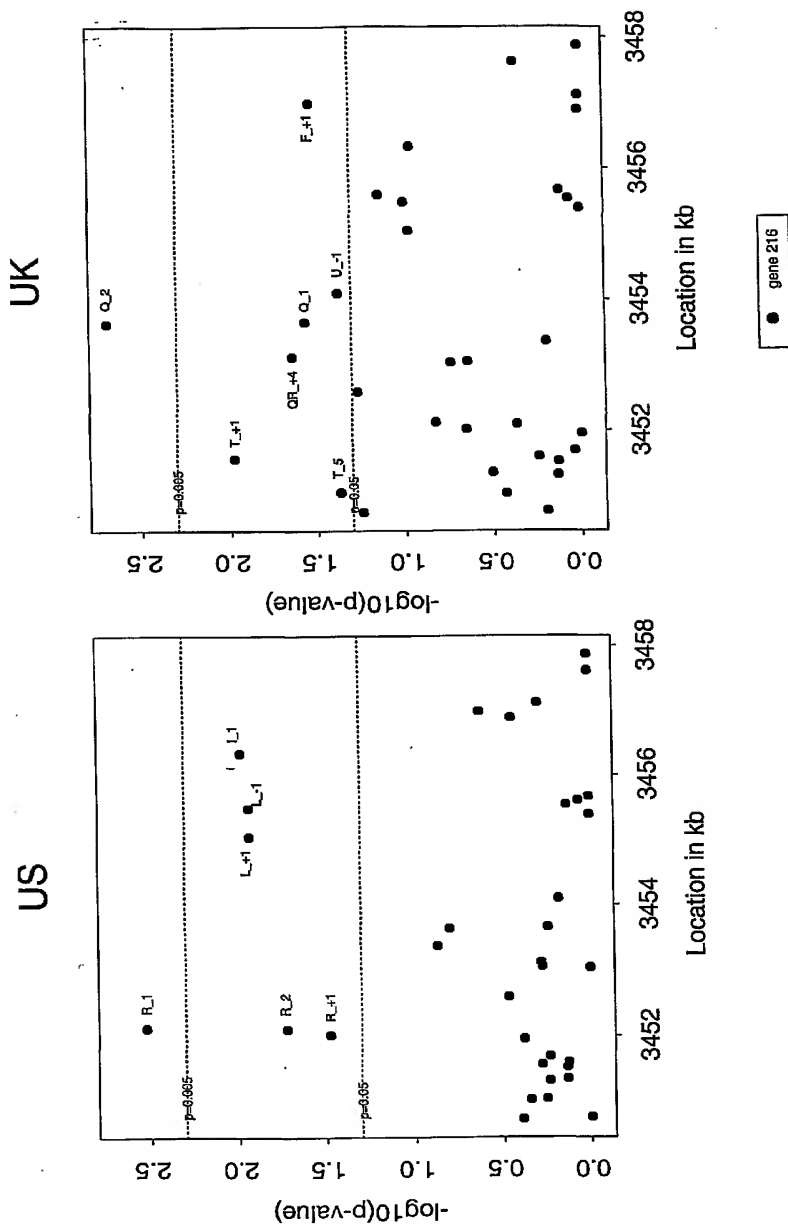


FIG. 26

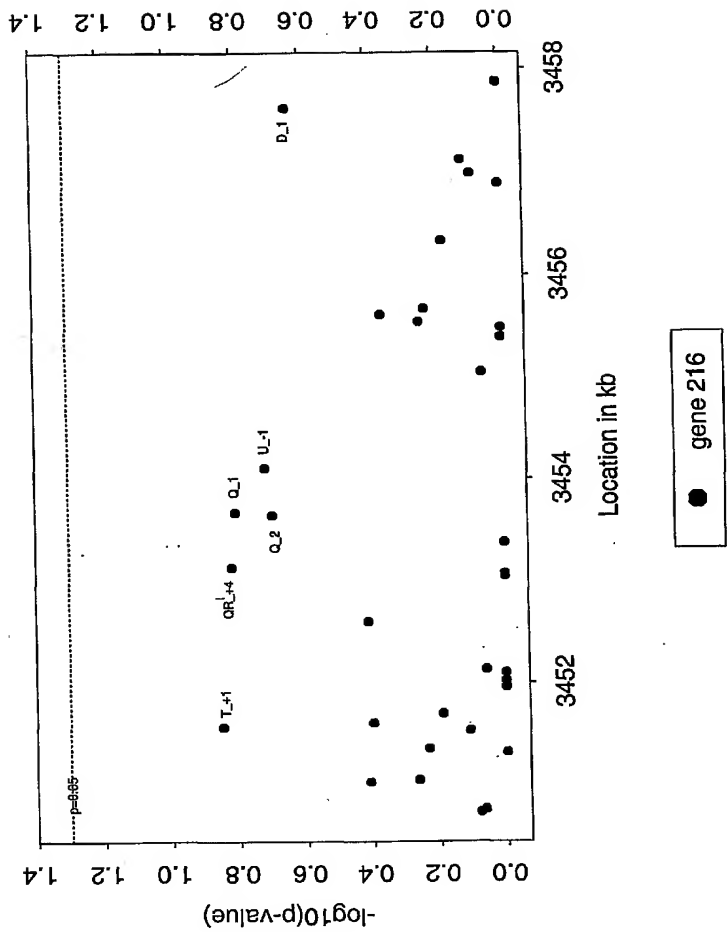


FIG. 27

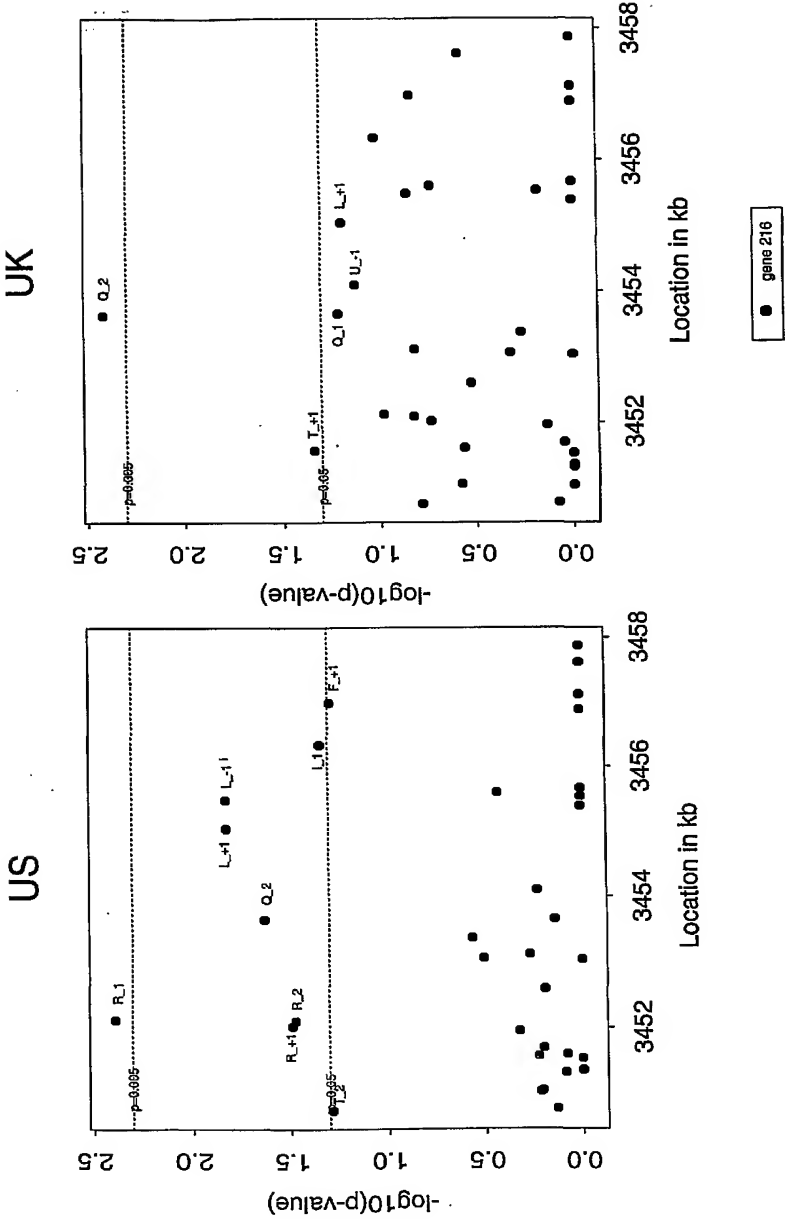


FIG. 28

94/157

```

      10              30              50
CGCAACTGGAGTGAGTCTCTTTCTTGCAACCCAAACAGCTGGGCTAGAACAGAGGAGTCA
-----+-----+-----+-----+-----+-----+-----+-----+
      70              90             110
CATGTTAGTGGGGTTTATTCTTAGACACCATCTGTAGGAGGGGGCCAGGGGCTGCAAGTT
-----+-----+-----+-----+-----+-----+-----+
      130             150             170
ACATGACATTATCACAAATTACCTTACGCTGAAGTGATTGTGTCCCTTTCTGCTCCTT
-----+-----+-----+-----+-----+-----+-----+
      190             210             230
GAGGCTTCGTGTGTGCTCACATTGCCACCCACAATGACCAGCCCCGTGCATGGCACGC
-----+-----+-----+-----+-----+-----+-----+
      250             270             290
AGTAGGTGGGCACGTGAACCTAAGCGTTTGTCTTGTCTACATTGCTGACAGTGGGGAGGGT
-----+-----+-----+-----+-----+-----+-----+
      310             330             350
GGAGAACATGGACTAAATCCAGAGGTGGGACAGGGAGAGAGCAAGTCAAATGCAGCCAGA
-----+-----+-----+-----+-----+-----+-----+
      370             390             410
TGTTTGGGGGCTGTCTGGGGAAGAGAGGGTGTAGGGAGGACCATAAGCAGCTGCTGAAGC
-----+-----+-----+-----+-----+-----+-----+
      430             450             470
TGCGGCTACTGGGGGATTGGGAATATCCGGCTTAGGAGTCTTCACCTGCTTCTGCATA
-----+-----+-----+-----+-----+-----+-----+
      490             510             530
CTTCCAAGGGGACATATGTCCATCTTGGATTGGAGGCTCTCAGAGGGGTAGTTGTGTCC
-----+-----+-----+-----+-----+-----+-----+
      550             570             590
```

FIG. 29

95/157

GCTTCCAGTTGTCGAGGCCTCCTGAGACTTTTTGTACATTCAATTAAGTGTAACATAC
-----+-----+-----+-----+-----+-----+

610 630 650
AGAAAAGTTCACACATTATAAATGCAACTCAAGGACTTTTCCAAAAGCGAACACACCCAG
-----+-----+-----+-----+-----+-----+

670 690 710
ATCAAGAAATAGACCATCCTACAGTCCCCCTTACACTCTGTACCAGTTGCAGCCCCAC
-----+-----+-----+-----+-----+-----+

730 750 770
AAGGGTAACTACTGTCTTGACTTCGAACACCATAGATTTGATTTGCCGTGTTTTAAACTT
-----+-----+-----+-----+-----+-----+

790 810 830
TACATAAGTAGAATCACAGAGTGTGTACAATGACTTTGGAAAAGTGTGTTGACAATATCTA
-----+-----+-----+-----+-----+-----+

850 870 890
TTAAAGCTAAAAATACCTTGCCCTATGAACCTGAAATTCACCCCACCTTGCCAAGGGACA
-----+-----+-----+-----+-----+-----+

910 930 950
AAAAGTTCCCTCTAAATGCACCAGGCTGTCAGGGATGAAGCGTTGGCTTTGGGGCCCCC
-----+-----+-----+-----+-----+-----+

970 990 1010
ATTACACACATGACCTTTTCTGGGGCACCCAAGCATCAGCCTGTCGTCACCAAGGTGCCA
-----+-----+-----+-----+-----+-----+

1030 1050 1070
CCCTGGCGATCTCTGAAGGCTGGAGTCGGAGTGCCTCCCTCAGACATCCTGTTCTGCGTC
-----+-----+-----+-----+-----+-----+

1090 1110 1130
ACTCCTTGGGAGAAGTCGTGTTTACAGATGGTGGGTGTCACCCATGCCAAGCACTTCTAA

FIG. 29

96/157

1150 1170 1190
 GGGTTAATGCTCACTGGTTTGCCTGGTTCCCGAGGACATTTCCTGATGCCCTCTGGAGGG
 -----+-----+-----+-----+-----+
 1210 1230 1250
 TGACGCCAACAAAGCCAGTGAGAGAACCATCTTCCCAGGTGCTGTCAAGCGCCCCCGGAG
 -----+-----+-----+-----+-----+
 1270 1290 1310
 CTGCTCGGTGCATCCTAGGATCCCTCTTCCTCAGCTTTGGTTTGATGGCTCATCTCCTC
 -----+-----+-----+-----+-----+
 1330 1350 1370
 CCCTGCAACCTCAAATGTAAATAAACCCCTTTCTCAGAGACTTCGGCAGAAAATTCCTCT
 -----+-----+-----+-----+-----+
 1390 1410 1430
 GACCTGCACTTGGACACAGCTCATCTGGGTTTGGGAGGTGTCAACTGTGTAAGGATGACT
 -----+-----+-----+-----+-----+
 1450 1470 1490
 CTGATCCCCATGTGGCTTTTCGACTGTGTCCCCTCTACAGTCAGTTATTAGCACTGACTG
 -----+-----+-----+-----+-----+
 1510 1530 1550
 TGCTAGGAAGTGAGCAACACACATATTCCCAGACCACATGGAGCTCAGGAGCTTGGGGAG
 -----+-----+-----+-----+-----+
 1570 1590 1610
 AGAGACAGGGAAGTGACGACTACAGGGCCTTCTGAAACGTGTTGCAGGGAGAAGTGTC
 -----+-----+-----+-----+-----+
 1630 1650 1670
 GTCAGGGGATGCTAACCTGGCTTTGGGTAAGGGACAGCCTCTGAATGACAGGACATTAA
 -----+-----+-----+-----+-----+
 1690 1710 1730

FIG. 29

```

GCCATGGCCTGCAGTTTAAGTAGGAGTTGGCCAGTTCGAGGTAAGAATACCAAGTAAGCAA
-----+-----+-----+-----+-----+-----+
          1750                      1770                      1790
GAACGCCAGAGTAGCTCCTCGAGCTGCCTTCTGTACCTGACATCCACACTGAAGCCAGCC
-----+-----+-----+-----+-----+
          1810                      1830                      1850
CCTCTGTGTTTCAGCCTTGCTTTACTGAAGAGGTGTCGCTGAGGGGCTGCTCTGGGGTGCT
-----+-----+-----+-----+-----+
          1870                      1890                      1910
GCTCTGCTTTCTCTGTCCCCAACTTGTTCTGAGCTCGAGCCACCTCCATACTGGTGTCTCT
-----+-----+-----+-----+-----+
          1930                      1950                      1970
GGTTCTCAGGCCTTTGAACTCAAACCTGAATCACACCCTGCGCTTTCCTCGTTCTCCAGCT
-----+-----+-----+-----+-----+
          1990                      2010                      2030
TGCAGATGGCAGATTTCGGGAACTTTTGGCCTCCATAATCACGTGAGCCAATTGCTATAA
-----+-----+-----+-----+-----+
          2050                      2070                      2090
TAAATATCTCTCTCCCTCTTTCTTCCTCTCTCTCTCTCTGTGCAAAATATAGTTCCAATTA
-----+-----+-----+-----+-----+
          2110                      2130                      2150
TAAGAGCCCCCTAACTGGAATAAACCCTATGGTGCACCTGGTGAGTAGAGAAACTGTGGTT
-----+-----+-----+-----+-----+
          2170                      2190                      2210
CCCTCAAACCACCGAACACTATTTCAGCAATACGAAGGAACAAACTATTGATATGCAAAAT
-----+-----+-----+-----+-----+
          2230                      2250                      2270
AGTGTAATGAATCTCAAAAACATCGGAAAGAGGGAAGGAAGCCAGACACAGAAGAGTGTC
-----+-----+-----+-----+-----+

```

FIG. 29

98/157

```

2290                2310                2330
ATGCCGCATGATTCCATTTATATGAAATTCTAGAACAGGC AAAACTTATCTATAGACAGA
-----+-----+-----+-----+-----+-----+-----+

2350                2370                2390
GAACAAACAGATCAGTGGCTGTCTGGGGTTGGGAGTGGGGAAGTTTGGCTGGAAGGGCACAA
-----+-----+-----+-----+-----+-----+-----+

2410                2430                2450
GGGCTCTTTCTGTGAGTGAGGGAATGTGTCTGCATTATAGTGATGCTTATGTAGTTATAT
-----+-----+-----+-----+-----+-----+-----+

2470                2490                2510
ACACTTATCGAAACTCATCTTACTGGCCACTTAAAAATAAGTGCATTTTATTGTGTGTAAA
-----+-----+-----+-----+-----+-----+-----+

2530                2550                2570
TTATACCTTAATGAAGTTGATTTGAAAATCCAAAGTAGTAATAATAAGTAATAATCTCGT
-----+-----+-----+-----+-----+-----+-----+

2590                2610                2630
AGCTGGACAGCTGTGGTGACTCACTCCTGTAATTCCAGCGATTTGAGAAGCTGAGGCAGG
-----+-----+-----+-----+-----+-----+-----+

2650                2670                2690
AGGATCACTTAAGATCAGGAGTTCTTTTATTTTATTTTATTTTATTTTGTGACGGAGTTT
-----+-----+-----+-----+-----+-----+-----+

2710                2730                2750
CGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCATGATCTCGGCTCGCTGCAACCTCCACC
-----+-----+-----+-----+-----+-----+-----+

2770                2790                2810
TTCTGAGTTCAAGCGATTTTCCTGCCTCAGCCTCCCAAGTAGCTGGAACACTACAGGCGCTC
-----+-----+-----+-----+-----+-----+-----+

2830                2850                2870

```

FIG. 29

99/157

```
ACCACCATGCCCCGGCTAATTTTGTATTTTGTAGAGATGGGGTTTCACCATGTTGGCC
-----+-----+-----+-----+-----+-----+-----+
2890                2910                2930
AGACTGGTCTTGAACTCCTGACCTCCAGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTG
-----+-----+-----+-----+-----+-----+
2950                2970                2990
GGATTACAGGCATGAGACACTGCGCCTGGCCAAGACCAGGAGTTTGAGACCAGCCTGGGA
-----+-----+-----+-----+-----+-----+
3010                3030                3050
AACAAAGTGAGACCCCTGTCTACAGAAAAATTAAAAATTTAGCTGGGCCTGGTGCCGT
-----+-----+-----+-----+-----+-----+
3070                3090                3110
GTGCCTGTAGTTCCAGCTACTCAGGAGGCTGAGGTGGGAGGATACCTTGAGCCCAGGATT
-----+-----+-----+-----+-----+-----+
3130                3150                3170
TCAAGGCTGCAATGAGGCATGATCAGGCCACTGTCCTCTAGCGTGGGTGACAGAGTGAGA
-----+-----+-----+-----+-----+-----+
3190                3210                3230
CCCTGTCTCTAAATAATAATCATAAGAACAACAAGGACCCTCTAAACGCACTGATATCTA
-----+-----+-----+-----+-----+-----+
3250                3270                3290
AGGTGTATTAAAGCGACCAAAAAAAAAAAGAAAATCAAAGTGCAGAAAAACGTTAATAAGA
-----+-----+-----+-----+-----+-----+
3310                3330                3350
GAAAAAATATGTCGTATTGTCTTGAGTGTGAAAAAATAATCTAAAGCCTATGAAAGA
-----+-----+-----+-----+-----+-----+
3370                3390                3410
AACTAATCATATTGGTTTCCTGTGTTGGTGAGGAGGGCTAAGAGCAGCGAGACTTTTCCCTA
-----+-----+-----+-----+-----+-----+
```

FIG. 29

100/157

```
      3430              3450              3470
TGCTTTCTGTACTTTTGTATTTTGAGATATGTGAATGTAGGTTTCTCTCACTGCTCGAAC
-----+-----+-----+-----+-----+-----+-----+-----+
      3490              3510              3530
TTTCACTAACCAAATTACTACATTCCAAATTCTCAAAAACAAATAGATTTACTTAAAGT
-----+-----+-----+-----+-----+-----+-----+-----+
      3550              3570              3590
AGGCTGGGTGCGGTGTCTCACGCCCTGTAATTCCAGCGCTTTGGGAGGCCGAGGCGGGCAG
-----+-----+-----+-----+-----+-----+-----+-----+
      3610              3630              3650
ATCACCTGAGGTTCGGGAGTTCGAGACCAGCCTGACCAACATGGAGAAACCCCATCTCTAC
-----+-----+-----+-----+-----+-----+-----+-----+
      3670              3690              3710
TAAAAATACAAAATTAGCCAGGCGTGGTGGCGAATGCCTGTAATGCCAGCTACTCGGGAG
-----+-----+-----+-----+-----+-----+-----+-----+
      3730              3750              3770
GCTGAGGCAGAAGAATCACTTGAATCTGGGAGGCAGAGGTTGCAGTGAGCCCAGATCATG
-----+-----+-----+-----+-----+-----+-----+-----+
      3790              3810              3830
CCATTGCACTCCAGTCTGGGTAACAAGAGAGAAACTCTGTCTCAAAAAAAAAAAAAAAAAA
-----+-----+-----+-----+-----+-----+-----+-----+
      3850              3870              3890
AAAAGATTGTCTTAAAGTTAACATCTCCGGCCGGGCGCGGTGGCTCATGCCTGTAATCC
-----+-----+-----+-----+-----+-----+-----+-----+
      3910              3930              3950
CAGCGCTTTGAGAGGCCGAGGCGGGTGGATCACGAGATCAGGAGATTGAGACCATCCTGG
-----+-----+-----+-----+-----+-----+-----+-----+
      3970              3990              4010
```

FIG. 29

```

CCAAAATGGTGAAACCTCGTCTCTGCTAAAAAATACAAAAGTTAGCTGGGGGTGGTAGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+
          4030                      4050                      4070
CGGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCAGGGAGT
-----+-----+-----+-----+-----+-----+-----+-----+
          4090                      4110                      4130
CGGAGGTTGCAGTGAGCCAAGATCGCGCCGCTGCACTCCAGCCTGGCGACAGAGGGAGAC
-----+-----+-----+-----+-----+-----+-----+-----+
          4150                      4170                      4190
TCCATCTCAAATAAAAAAAAAAAAAAAAAAAGTTAACATCTCATCCAAATTGCACCGAGTA
-----+-----+-----+-----+-----+-----+-----+-----+
          4210                      4230                      4250
GGAAACAAAAGTTTAAACATGAAACAGATGTTACTGAGGCCGAAAGGGTCTCCAGGC
-----+-----+-----+-----+-----+-----+-----+-----+
          4270                      4290                      4310
CTGGGAGTCTGCAGCTTTTATGCAATTCTGCCCTCTGGCCACCGCCAGGGAAGAAAGGTT
-----+-----+-----+-----+-----+-----+-----+-----+
          4330                      4350                      4370
GTCTCCGTCTGCTGCATCGCCTTTGCCCAGCAATGAAGCCCCCAAGACAGCGGCAGCCGG
-----+-----+-----+-----+-----+-----+-----+-----+
          4390                      4410                      4430
TTGCCTGAACCTTCCTATCCTTGGGGGCACCCAGTGCAAGGTGGATGACCCGACTCAACCT
-----+-----+-----+-----+-----+-----+-----+-----+
          4450                      4470                      4490
CCGCCAGGGCACCTCTGGGGCAGGACGGGTAGCAAGGAGGGGACAGAGATCGSCCCCAGG
-----+-----+-----+-----+-----+-----+-----+-----+
          4510                      4530                      4550
AGACCACGGAAGATCGCGCTCCTGGGGCCAACTTCAGCAGCGAGAGCGCGCCTTTGCCCA

```

FIG. 29

102/157

```

      4570              4590              4610
CCGCTCATCCACACGCGCGGTCCTCCAAGAACCTTCCAGCGGTTCTCTCTCTCTC
-----+-----+-----+-----+-----+-----+-----+-----+

      4630              4650              4670
TCAGGAGTAGAGGCCCTCTGAGACCGACGGGGAGGGACGGCTCGGGCCGGTCATCCGAGG
-----+-----+-----+-----+-----+-----+-----+-----+

      4690              4710              4730
GGCCGCACGGGATCCCTCTCTCCGCCAGCTCCACCCCTCGAGGGGGCGGGCTCCGGGAG
-----+-----+-----+-----+-----+-----+-----+-----+

      4750              4770              4790
TGGCGACCCGGCTCCCCATGGCGCGCGCCGTCGGGGCCCTGGCCAGGCTCCGAGCGGG
-----+-----+-----+-----+-----+-----+-----+-----+

      4810              4830              4850
GTTGGCGGGGAGGGGAGGCGGGAGCGAGGGCGGGCGGTGGGAGGTGGGGCGGGAAGGTC
-----+-----+-----+-----+-----+-----+-----+-----+

      4870              4890              4910
CGAAGGCGGCGGCTGAGGCTGCACCGGGCACGGGTCGGCCGCAATCCAGCCTGGGCGGA
-----+-----+-----+-----+-----+-----+-----+-----+

      4930              4950              4970
GCCGGAGTTGCGAGCCGCTGCCTAGAGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCC
-----+-----+-----+-----+-----+-----+-----+-----+

      4990              5010              5030
CGGAGAGCTCGGGGACCCCGTTGCTGCTGCTGCTACTACTGCTGCTCTGGCCAGTG
-----+-----+-----+-----+-----+-----+-----+-----+

      5050              5070              5090
CCAGGCGCCGGGGTGCTTCAAGGTGAGGACGCGGGCGGGGTGCGCCCTGAGGGGCAGGCT
-----+-----+-----+-----+-----+-----+-----+-----+

      5110              5130              5150

```

FIG. 29

103/157

AGGCGCGGTGGTGGTCGGGGATGGGTCTTGCTCAGAGCTCGGGTCAGCGCGGAGGG
-----+-----+-----+-----+-----+-----

 5170 5190 5210
TCTCACGGCCCCGGGCACCATACGGCCAGTAGGTCAAGGCGTGGGGACTCTTTGGGGGGGT
-----+-----+-----+-----+-----+-----

 5230 5250 5270
CTCCGTGGGACCTGCCCGAGGACGCTCAAGTGTGCTTTGGGCTGGCCCCGGGCCCAGACTT
-----+-----+-----+-----+-----+-----

 5290 5310 5330
GCCCCACACTGCCCGGCTGCCACTCCGCTGGCAAAGCAGAGGGCATGGCTCCCTCCCCCTC
-----+-----+-----+-----+-----+-----

 5350 5370 5390
GGGGACAGCCCGAGCCCCAGCCCCAGCCCCATAGCCGTAGCCCCCTCTGCCTGGATTCTC
-----+-----+-----+-----+-----+-----

 5410 5430 5450
GCTCTCACAACCAGCTTCCATCCGCAGGCCACCGTGTGACCCGCTCCTGCTCCTCCACCC
-----+-----+-----+-----+-----+-----

 5470 5490 5510
CTTAGGACTCAGCGGGGCTCCATCTCTAGGAAGCCCCCATGCCAAGAGTCCCCAGAG
-----+-----+-----+-----+-----+-----

 5530 5550 5570
TCCCTGCTTTGCTCTCAGGCTGCAGAATACTGCTGTGGCCTCCACCCTGCTCACCCCTCGT
-----+-----+-----+-----+-----+-----

 5590 5610 5630
CCCTCTCTCCAGGGCAGCAGGGCAGTGTGTATGTTGTTTATATTGTTGCCTTGTTTGGTG
-----+-----+-----+-----+-----+-----

 5650 5670 5690
AGATAGAGAAGGGCCTCTCCAGATAGAAGGTGTCTGTTTAGCAGTGTCTCTGGAAAGACTG

FIG. 29

104/157

```
5710          5730          5750
CAGCTGTCCTCTCGGGTAACCCCTCCAAACAAAGATGTTAAGATGGGGCTGGAACAACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

5770          5790          5810
TCTGCAAGCGGGTGGGAGGATTAGCCAGTCCTGCACAGCAAGTGCCTGGCCGGGAACAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

5830          5850          5870
GAGGGCAACCAGGGAGGGGGCATGCGGGGCTGGGCTGTGCTATGCAGACTGGGCGGTGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

5890          5910          5930
TTCCACAGCACTGTGTGGGGACCAAACAGGTACAGGGGCCTGGTCTGTTCTGGCCCCAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

5950          5970          5990
GGAGGGCCCCAGGCGGTCCACTGCTCCCTCCCCTCTGAGCCCTATCCTGGGGTCAGGGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

6010          6030          6050
GGTGATGGGACCCCTGGGAGAGGGGCGTCTATGTGCCCAATACCAGCCTGGCTCCCTCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

6070          6090          6110
GTTCACCCCCATTACCCCGGTACCCGGAGCTCCAGCTCCAGCTCCAGCTCTGCCCTCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

6130          6150          6170
CTCCCTCATTTGGGGTCAGGGTGCCCGTGGCCAGCACGTGCGCGCAAGGCCATGTGGACAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

6190          6210          6230
CACCCACACACCACACTGCACCCACACCACACCTGTGCCGGGGCCACCCTACCTCTTCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

6250          6270          6290
```

FIG. 29

106/157

```
6850          6870          6890
CGTCCCTCCCTTGAAGAGATGAGGAGGGGCCTTCTGGGCCAGGGTACCAAAACCCCTGC
-----+-----+-----+-----+-----+-----+-----+
6910          6930          6950
CACCAGGACAGAGTCCCCGAGGAGCTCTGGGCAAGGTGGACCTCGCAAGGCAACATCTG
-----+-----+-----+-----+-----+-----+-----+
6970          6990          7010
GCTGTTGTTTTTCTCAGATGATGGGGGGGGCACAAGTGTCCTCTCTTCGTACATCTCTCA
-----+-----+-----+-----+-----+-----+-----+
7030          7050          7070
CCCTAAAGGCATCTGCTGCCCATCTAAAAATCCCTAAGGCTGCCGCGCTCTTTCCTTCCC
-----+-----+-----+-----+-----+-----+-----+
7090          7110          7130
CTCTGCACTGGCGGCCTTGGCCTCTTCCCTTGTGATCGCCGAGCCCAAGCCTGCCCCCCGA
-----+-----+-----+-----+-----+-----+-----+
7150          7170          7190
CAAAGGTCAGGGGACTCCCGTGTCCCCAGCTGAGCTGTCCCTTTCCAGCCTTCTCTTTTC
-----+-----+-----+-----+-----+-----+-----+
7210          7230          7250
CTCCTCCTTGATAGCTCCTCAGATCCAAGGATGCCACGGGCGTCCCTCCTTCTCCAGGC
-----+-----+-----+-----+-----+-----+-----+
7270          7290          7310
TGAGCCCA CGCGTGTGAAGGTGAAGTCTGCCCCAAAAGGCCTCAGTGCCTCCCTGGGG
-----+-----+-----+-----+-----+-----+-----+
7330          7350          7370
ATGTCCTCTACCCCCCTCCTCTGCTTTGTCCCATGCCCTGTGTCTCTCAGGTCCCCCT
-----+-----+-----+-----+-----+-----+-----+
7390          7410          7430
```

FIG. 29

107/157

```

CACCTGTGCTCTGTCTTTACTCCAGGACATATCCCTGGGCAGCCAGTCACCCCGCACTG
-----+-----+-----+-----+-----+-----+-----+
          7450              7470              7490
GGTCCTGGATGGACAACCCCTGGCGCACCGTCAGCCTGGAGGAGCCGGTCACTGCCATGTC
-----+-----+-----+-----+-----+-----+-----+
          7510              7530              7550
TCCCCGCCCTCCACAGGGGCCCTGAACCTCCCAGCCCTTTTGTCTCTCCCTACATTACAG
-----+-----+-----+-----+-----+-----+-----+
          7570              7590              7610
CTTCTAGTTTGTCTGGGGTCCCCAGAACCACCAAGTCACTACTCCTATAGGCCCTGCGCT
-----+-----+-----+-----+-----+-----+-----+
          7630              7650              7670
CCCCTGCCCTCAAGTGGGCAGAAGAAGGCACTGGGGTTTGGACATCTGGATCTCGTGAG
-----+-----+-----+-----+-----+-----+-----+
          7690              7710              7730
CCCGCACACATGGAAGTCATTTTCAGCTTTCTCCACCCCACTCCCTCTTTCCCTCCCTCC
-----+-----+-----+-----+-----+-----+-----+
          7750              7770              7790
CTGGATGATCTGGGCCACCCCCACCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
-----+-----+-----+-----+-----+-----+-----+
          7810              7830              7850
CCTGAAGCTTTTCAGGAGCCTCTAAAAAAAAAAAAAAAAAAAAAAAAAGCACCAGGAGGAGG
-----+-----+-----+-----+-----+-----+-----+
          7870              7890              7910
CCTTTTGCAAGTTGACCAGAACATGTGACCCCTGTGGACACACTGCTGTCCCTCTCAGGG
-----+-----+-----+-----+-----+-----+-----+
          7930              7950              7970
CCCTGCCACGAAGGCCTGAACCTTCAGCCTCACTGGCTCCTGTGGAATCCACTTCTGGTA
-----+-----+-----+-----+-----+-----+-----+

```

FIG. 29

7990	8010	8030
TGGGGGGGGCAGTGCTCACTCTCCTGATGTCCCCCAGATGTAAGACCACCCCATGTGCTT -----+-----+-----+-----+-----+-----+		
8050	8070	8090
CTTCTGCAGGACGCTCTGCCCGAGCCTCTTCCCAATCCCCTCTTCACACGCTTCCAGAA -----+-----+-----+-----+-----+-----+		
8110	8130	8150
TAACCATGCCCCATCTGTTTGTGCCATAATATCTGTGCTGCAAAC TAAGAGGGCAGTAGC -----+-----+-----+-----+-----+-----+		
8170	8190	8210
CTTGATATGCTCATTTTTACAGAGGGGCAAACGGAAGCCCAGAGAGCTTGGGGAAATTGTC -----+-----+-----+-----+-----+-----+		
8230	8250	8270
CATGGTCACACAGCTCTTTAGGCTGGGAGCCTGAGACCCACTAAGGCTCTGAACGATTTTA -----+-----+-----+-----+-----+-----+		
8290	8310	8330
AACCATTGGCTACACCCCTGCCCTCCTAGAGAGCCCTCTTTGTTTGGAAATTTTCAGCCC -----+-----+-----+-----+-----+-----+		
8350	8370	8390
TACTGTCCA AATCCAGCAAGAGGGAAGGCAGGGGAGCATTGCCATGAAGGCTGAGAGGCC -----+-----+-----+-----+-----+-----+		
8410	8430	8450
CCCAGAGACCCAGCAGCTCCCAACCCAGGGCCCTCACTGGGATCCCCCTAGGCCCCATAAGG -----+-----+-----+-----+-----+-----+		
8470	8490	8510
CCCCCATTCCTACTGGTCAAGCACGGCACTGGCCTGAGCTTTGAGATTGCCCTCCCCATCC -----+-----+-----+-----+-----+-----+		
8530	8550	8570

FIG. 29

109/157

CCAGGAGGGGAAGGCTGGACACACTGGGGTCACTCTGCCTCTGGGCCCTCCCTGTCTGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
8590 8610 8630
CTGGCCTGGGCTGTGACCAAGAGGAGAGCCCCAAAGGGGCTCTGTCTCCCCCACCGETGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
8650 8670 8690
GCCCCCTGCCCCAGGAAGCCTGCCAAGATGGTACAGAAGAAAGAGTAGAGGCTAGGTATC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
8710 8730 8750
CCCTCCAAAAGGCAGGAAACACTCACATTTCAAGATGAGGGGTATATATCAAGGGGCAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
8770 8790 8810
GTACCAGGAGGGCAAGAGTAAAGATAGCAGGGGCTGCAGAGGAACAGGGACCTCGAGTAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
8830 8850 8870
GGCCTTTTTCGGGTGCAGACCTTTCCCCAATAAAGCAAGTGGCATTCCAGCCTCATGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
8890 8910 8930
CTCATGCTGGAGGCCTTGTGGGGCCTGTGGCCAGGGAGGCAAGGACCATCTGCTCCCCAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
8950 8970 8990
TTGCGAAGGAAGAACTCCCTCCAAAGACTCTGAGACCTTGGACAGGGCCCCAGGCCAGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
9010 9030 9050
GCATTTTGGAGAAAAGGAGTCGGGGGTAAACATTCCGAAGGCGCAGCAGCCTCCCAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
9070 9090 9110
AAGCTCCTGGGCCGGCTCCAACTCTGGGCCCCAGCCAGGCTGAGTGGACAAGGGGGAAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

FIG. 29

9130	9150	9170
TGGGGTGTTC	CCACAGGGTGGGAGACGCCAAGAGGGTGGGGGAAGGAGAGAGGGGCTGGCC	
9190	9210	9230
GTCCAAGCCAGCCTCCTGACACCTAGCTGAGAGCCAGTGTGCTCTCTTGGCTGGAATGCC		
9250	9270	9290
GTCCATGTTTACTTCGTGGGTCCAGTGAAGCAGGTGTCCGAGCCGAGGGGACGGGGGGCTG		
9310	9330	9350
CTGGAGGCCCAGGAAAACTTTGAAGAGGGAGCAGTTTGCCAAAATTGGAAGTGGAGGAG		
9370	9390	9410
TCAAAATTTGAATTCTATAGGAAATGAGCAGCAGCTCATTTGGAACCAAGCCTCAGGTAGC		
9430	9450	9470
AGAGGGCTCTGAGGAGGCCCTGACCATGGGTACCCGATGCCCCCAATAATGTCTCAGCACC		
9490	9510	9530
CCTCTGTCTTCCCCTGCTTTTGATGCCCTTCTGGGCATGAAAGAAGAGGGCGGGGCCAG		
9550	9570	9590
GGGAGGGGCACCTTTCTGGGACCTCTGGTCTCTAGGGAGGATGCTGCTGTGCCTGGCAGG		
9610	9630	9650
CTGTGCCAACGCCCTTCCAAGTGGCTGTGTGTCAGGACTGCAAACATCCTGAGTTTGGGAA		
9670	9690	9710

FIG. 29

[illegible]

FIG. 29

112/157

```

      10270                      10290                      10310
CAGGGACCAAGATGCGGTGGCTTTGGGAGGCCGAGGCGGCGGCTCACCTGAGGTCAGGA
-----+-----+-----+-----+-----+-----+-----+
      10330                      10350                      10370
GTTTGAGACCAGCCTTGCCAACATGGTGAACCCCGTCTCTACTAAAAATACAAAAAATT
-----+-----+-----+-----+-----+-----+-----+
      10390                      10410                      10430
AGCCAGGCGTGGTGGCGGCCTGTAATCCCAGCTACTTAGGGGGCTGAGACGGGAGAAT
-----+-----+-----+-----+-----+-----+-----+
      10450                      10470                      10490
CGCTTGAACCCCGGAGGTAGAGGTTGTGGTGAGCCAAAGATCACACCACTGCACTCCG
-----+-----+-----+-----+-----+-----+-----+
      10510                      10530                      10550
GCCTGGGCAAAGAGTGAGACTCCGTCTCAAAAAAGAGAAAAAAAAAGAAAAAGAAAAAA
-----+-----+-----+-----+-----+-----+-----+
      10570                      10590                      10610
AAAAGAAAGAAAAGAAAAAGAAAAGATGCAGTGGCTACACTTGGGGGCAGCAGTTTGT
-----+-----+-----+-----+-----+-----+-----+
      10630                      10650                      10670
CTGACCTGCCTGGAAGGTCTCCATCTACAGGGAGGGGAGCAGGGGGGAATGAATTGGAG
-----+-----+-----+-----+-----+-----+-----+
      10690                      10710                      10730
AGTCCCAGGAGGGGCCAGATCACAGAAGGCCATTTTGGTGCTCAGTGTCTCTGGACCATCA
-----+-----+-----+-----+-----+-----+-----+
      10750                      10770                      10790
GAGCCAAAGATTTTGAGCTTGGGGAAGGGACAGGCAGACCTGTGCTCAGGAAGGTGCCTTG
-----+-----+-----+-----+-----+-----+-----+
      10810                      10830                      10850

```

FIG. 29

113/157

GGCTGGGTGGGGTGGGTGTCGGGGCTGGAGCGCAGGCTCTTAAACCACCCAGATTATGT
-----+-----+-----+-----+-----+-----+

10870 10890 10910

TATCAGTATATATCACCTACTGAGTGCTTGACCGCAGGCCTGTTCTGAGCACTTGACAC
-----+-----+-----+-----+-----+-----+

10930 10950 10970

GTA TTTTATTTCTCCCTCGTGGAGTGGGATGGACAGGGAACAAACTCTAGTITCCACTGTGC
-----+-----+-----+-----+-----+-----+

10990 11010 11030

C CAACCATATTTTTCCCCGACGTCCTTACCTTTTCAATGGGGTGGTCACATCACCTACCTCC
-----+-----+-----+-----+-----+-----+

11050 11070 11090

TAGGGTGGCGGGTG GTGTGTGGGGCAGGGGTAGGGGGCAGAGCTGGGGCAGGTGGTGGAAAT
-----+-----+-----+-----+-----+-----+

11110 11130 11150

G CCTGGGAGGGGGGAAGCAGCCATCATTAGCGGGTGGTCTGGAGGTAATGAGGCCAAGGT
-----+-----+-----+-----+-----+-----+

11170 11190 11210

G AGGTTGGGTTAAGGATTTTCTTTAAGAAGACAGATTGACTTATGATTGATCCATCCGT
-----+-----+-----+-----+-----+-----+

11230 11250 11270

G TGGGAAAGATCCTGTTGAGATGGAGCCTGAAGATGGAATCATTACCGGAGTGGGTGTGG
-----+-----+-----+-----+-----+-----+

11290 11310 11330

A GAAGGCAGGGAGGGTGGGAAGCAGCGTGGGCAGGTGGCGATTCTGTTTTCTCTGGAGGCA
-----+-----+-----+-----+-----+-----+

11350 11370 11390

G GGGGTTGAGCATCAATCACTGAAGGACAGGTGGGAGGTATGTGGGGTCTAGAAGTCTGAG
-----+-----+-----+-----+-----+-----+

FIG. 29

114/157

```
11410      11430      11450
GAAATATTTCAAGGATCTAGGGCAGGTGGGGGCAAGAGGGTCGACCAGATGCCCAACAA
-----+-----+-----+-----+-----+-----+-----+
11470      11490      11510
AGGAGGGCAGCAGGCAGGGGAAC TGGGGGAGGTCACCGCATTTCCCAACTCCAAGTCCC
-----+-----+-----+-----+-----+-----+-----+
11530      11550      11570
ATTCTTCGGCAGTGTCTCCTGACTCCTCCCCTCCCGATCCTGTGGATCCTGCTGCCTGCT
-----+-----+-----+-----+-----+-----+-----+
11590      11610      11630
GCAGGTCCCCTGGGAACCACAAACTCTTCCCCTATTCCCACTCCTCCCCGGCGTCTCCC
-----+-----+-----+-----+-----+-----+-----+
11650      11670      11690
TGGTGCTTCCCATATTCACATCTCCCACTAAGCCATCACCAAGGCTCCTTCCTCTAG
-----+-----+-----+-----+-----+-----+-----+
11710      11730      11750
CCCCAAGAGTTTCTGATCTGAGCAAGTCACCATTGCTCCTGTCCCTTCCCTAAGACACAC
-----+-----+-----+-----+-----+-----+-----+
11770      11790      11810
TGTGAGTGTCTCACTCATAAAGCTGCTCCATTAGCATTAGGGAGGAAGGCTGGGAGACA
-----+-----+-----+-----+-----+-----+-----+
11830      11850      11870
TCCTGGAGGAGGCAGGAGGAAGCTGAATTCAGTGTTCCTGTAAACCCCCCTCTCAGCAG
-----+-----+-----+-----+-----+-----+-----+
11890      11910      11930
GCTGCTGGCCCCAGGATACATAGAAACCCACTACGGCCCAGATGGGCAGCCAGTGGTGCT
-----+-----+-----+-----+-----+-----+-----+
11950      11970      11990
```

FIG. 29

115/157

GGCCCCAACACACGGTGAGATGCTTCATGGGCTCTGGGATGCACCGCCAGAGGTACC
-----+-----+-----+-----+-----+

12010 12030 12050

CCCCCACCATTCTACCCCTACTCCTCCTGCAATTCCTAAGGGGCGGTGGAGCCAGCCC
-----+-----+-----+-----+-----+

12070 12090 12110

CTACCACACCCCTCCTTGC CCCCTCTTGCTCCAGCCCTGGCTGAGATTGGGGCTGGCC

+-----+-----+-----+-----+

12130 12150 12170

CCTTCCTCCCTAGGATCATTTGCCACTACCAAGGGCGAGTAAGGGGCTTCCC CGACTCCTG

+-----+-----+-----+-----+-----+

CGTCTG

12190 12210 12230

G**TAGTCCTC**TGCACCTGCTCTGGGATGAGGTGAGCTCTGGGAGAGGAGGCTGGGCCCTGG
+-----+

12250 12270 12290
GATGGGGAAAGAGCTCCCTCACCCCGCTCCTACCCCTCTGCACCCTAGTGGCCTGATCA
+-----+-----+-----+-----+-----+-----+-----+-----+

CCCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGGCCACCCGGGGCTCCAAGGACT

TCTCAACCCACGAGATCTTTGGGATGGAGCAGCTGCTCACCTGGAAAGGAACCTGTGTGCC

ACAGGGATCCTCGGAACAAGCGGCGCATGACCAGCCCTTCCTGGTGGTCCCCAGAGCAGGG

TCAGGGGCTCGATCGGATGGGAGTGGGAATGCTGTATCTATAGCCCTCCAATCAGAAG

FIG. 29

116/157

```
12550      12570      12590
AGACAGGAATTTCACAGGCCCTCGAGTCCAGTATTTTATTGAAGTCTGAAGAAACAAGTT
-----+-----+-----+-----+-----+-----+-----+
12610      12630      12650
CCAGAAAACATGTTAAACTTCCTTCTGGGAGCTGGGGTTGGGGGTCAAGGCCCA
-----+-----+-----+-----+-----+-----+
12670      12690      12710
GCAGCTTCCACTCAGGGTCCCCATTTCACCTCCGCAGGGCAGGCGAGAAGCGCGCAGGA
-----+-----+-----+-----+-----+-----+
12730      12750      12770
CCCGGAAGTACCTGGAACGTGTACATTGTGGCAGACCACACCCTGGTGAGGAGAGACCCCA
-----+-----+-----+-----+-----+-----+
12790      12810      12830
GGGGTTGGCGGGGTCAAGGATGGGGCCAGCTCAGCCCCCTCAAGCCACCGGGATTCTCGCC
-----+-----+-----+-----+-----+-----+
12850      12870      12890
TTCCCAGTTCTTGACTCGGCACCGAAACTTGAACCACACCAAACAGCGTCTCCTGGAAGT
-----+-----+-----+-----+-----+-----+
12910      12930      12950
CGCCAACTACGTGGACCAGGTTGGGGGCGGCGGGGAGAGAGCGGTGATGGGGGTGGCGGC
-----+-----+-----+-----+-----+-----+
12970      12990      13010
GGCAGGACAGGCAGGTGCTGGTGGGGTTTGGGGGAAGAGGAAGGGCGCCCCACGAAGGACC
-----+-----+-----+-----+-----+-----+
13030      13050      13070
ACCGGCGCGATGGGGCGCCCTGTCCCGGCTTCAGCCCCGCCTCGCCCTCAGCTTCTCAGG
-----+-----+-----+-----+-----+-----+
13090      13110      13130
```

FIG. 29

ACTCTGGACATTCAGGTGGCGCTGACCGGCCTGGAGGTGTGGACCGAGCGGGACCGCAGC
-----+-----+-----+-----+-----+-----+-----+

CGCGTACAGCAGGACGCCAACGCCACGCCTCTGGGCCCTTCCTGCAGTGCGCCGGGGCTG

13210 13230 13250
TGGGCGCAGCGGGCCCCACGACTCCGCGCAGCTGCTCACGTGGGTGCCTCTGACCCGGACG

CGGGTCCC GGGTGGG GCGGCCTCACCTCC CGCCCCG CCTGGTCA GCAGC CGCCTCCGCC

13330 13350 13370
CCAGGGGCGCGCCTTCACAGGGCGCCACAGTGGGCCTGGCGCCCGTCGAGGGCATGTGCC

GCGCCGAGAGCTCGGGAGGCCTGAGCACGGTGAGCCCCGCCGGGGCGAGGGAGAGA

13450 **13470** **13490**

CAGGAGGCTCTACGGCCCGAGTGACCGCCCTCCCACGGCCCCCAGGAACCACTCGGAGCT
-----+-----+-----+-----+

13510 13530 13550
 CCCCATCGGC⁺CGCGCAGCC⁺ACCATGG⁺CCCATGAGATCGGCCACAGCCTCGGCCTCAGCC⁺A
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

13570 13590 13610
CGACCCCGAGGGCTGCTGCGTGGAGGCTGCGGCCGAGTCCGGAGGCTGCGCTCATGGCTGC

13630 13650 13670

GGCCACCGGGTACGCGGGTGGGGGGTCGGGGCTGCGGGGGGCGGCTAGTCTCTGGGGACT

FIG. 29

118/157

```
13690      13710      13730
TCCTCCGCTGCGTTTCTTTGGTCGTCCTCAGTTTCCTCTTCTGTAAATGGGGATAATG
-----+-----+-----+-----+-----+-----+-----+-----+

13750      13770      13790
ATCATAGTGTCCGCTTCAGGGTGCTTTATGAGGCTTAAAGGGAAGAAGCTCAGGCAAAGT
-----+-----+-----+-----+-----+-----+-----+-----+

13810      13830      13850
GGATTCTCAACGGTATGAAGATTATTTCCGAGTAACCTGGCGAGGTTACTCCTACACCG
-----+-----+-----+-----+-----+-----+-----+-----+

13870      13890      13910
GGAGGAGCACCGTCGGGTCGCGATTCCACCTTGGGTCCCGGGCTGCTCACTATTGGGGCC
-----+-----+-----+-----+-----+-----+-----+-----+

13930      13950      13970
GCATCGCTCCCCTGTCCCCTTGTGTGTGACTTTGCGCGGGTTACTTCCCCTCTCTGGGC
-----+-----+-----+-----+-----+-----+-----+-----+

13990      14010      14030
TCTGCGCGTCTGCGCGCTGTAGCCAAGCCCAGGGGTGGGGATCAGAGAAGCGCGGGGGTT
-----+-----+-----+-----+-----+-----+-----+-----+

14050      14070      14090
GGGGGACTGTCCCTCCATGCCCAATGCCCTCCCCGTGCCGCTAGGCACCCGTTTCCGCGC
-----+-----+-----+-----+-----+-----+-----+-----+

14110      14130      14150
GTGTTCAGCGCTGCGCCGCCGCCAGCTGCGCGCTTCTTCCGCAAGGGGGGCGCGCT
-----+-----+-----+-----+-----+-----+-----+-----+

14170      14190      14210
TGCTCTCCAATGCCCCGACCCCGGACTCCCGGTGCCGCCGGCGCTCTGCGGGAACGGC
-----+-----+-----+-----+-----+-----+-----+-----+

14230      14250      14270
```

FIG. 29

119/157

TTCGTGGAAGCGGGCGAGGAGTGTGACTGCGGCCCTGGCCAGGTTAAGTCGGCTCGCCCC
-----+-----+-----+-----+-----+

14290

14310

14330

GCCCCCACTTGCCCTCTCCGCTCAGGTCTGGGGCGCTGCGCCCTCACCTGGGCCCTTCTT
-----+-----+-----+-----+-----+

14350

14370

14390

GCCTTTCTGGTCCCAGGAGTGCCGCGACCTCTGCTGCTTTGCTCACAACTGCTCGCTGCG
-----+-----+-----+-----+-----+

14410

14430

14450

CCCGGGGGCCCACTGCGCCACGGGGACTGCTGCGTGCGCTGCCTGGTGAGGGCATGGAA
-----+-----+-----+-----+-----+

14470

14490

14510

GGTTCAGGGTGAGGGTTTCGGGGAGCTTGGGAGCCGGCCTGTTGGCCTTAGTTAATTGGT
-----+-----+-----+-----+-----+

14530

14550

14570

GCCCTCAGGTTCCCCCGTTGGGTGCTGGGCTTGGGTAGGCCTGGCTCCCCCAGCTCCGAG
-----+-----+-----+-----+-----+

14590

14610

14630

CCGCGCTCTCGGCATGGACCTCTCACTGCACGTGGCCTCTCTCTGCGCTTCCCCACCACCC
-----+-----+-----+-----+-----+

14650

14670

14690

GTCACCTGCGCAGCTGAAGCCGGCTGGAGCGCTGTGCCGCCAGGCCATGGGTGACTGTGA
-----+-----+-----+-----+-----+

14710

14730

14750

CCTCCCTGAGTTTTCACGGGCACCTCCTCCCACGTGCCCCAGACGTTTACCTACTGGA
-----+-----+-----+-----+-----+

14770

14790

14810

CGGCTCACCTGTGCCAGGGCAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGA
-----+-----+-----+-----+-----+

FIG. 29

14830	14850	14870
GCAGCAGTGC	CAGAGCTCTG	GGGGCCCTGGT
GAGAGGACAC	GAGCACCCCT	TGCACCCCTGC
-----+	-----+	-----+
14890	14910	14930
CCCCATCCTCT	GGTGGGGCCAGT	TTTCTACTGTGGGGA
AGATGGGCAGG	GGGAAACTGAG	
-----+	-----+	-----+
14950	14970	14990
GCCCGCTGAGCG	CAGCCCCCTCTCC	GAGCTGCCCCAGCCT
TGGCCCATGCTT	CCTCAGGCT	
-----+	-----+	-----+
15010	15030	15050
CCCACCCAGCTCC	CGAGGCCTGTTTCC	AGGTGGTGAAC
TCTGCGGGAGAT	GCTCATGGAA	
-----+	-----+	-----+
15070	15090	15110
ACTGCGGCCAGGAC	AGCGAGGGCCACTT	CCTGCCCCTGTG
CAGGAGGTAGGAG	TGGAGC	
-----+	-----+	-----+
15130	15150	15170
TGAGTGGAGGGAGC	AGAAGCTATGGAGT	TGGGTTTGGGGA
AGGGGTACTGCAGC	TGTTG	
-----+	-----+	-----+
15190	15210	15230
ACCCCCCTCTACTT	CCTCCCCAGGGATG	CCCTGTGTGGGA
AGCTGCAGTGCAGG	GTGA	
-----+	-----+	-----+
15250	15270	15290
AAGCCCAGCCTGCT	CGCACCGCACATG	GTGCCAGTGGACT
TACCGTTACCTAGAT	GGC	
-----+	-----+	-----+
15310	15330	15350
CAGGAAGTGACTTGT	CGGGAGCCTTGGC	ACTCCCCAGTCCC
AGCTTGGC		
-----+	-----+	-----+
15370	15390	15410

FIG. 29

121/157

CTGGGCCTGCTAGAGCCAGGCACCCAGTGTGGACCTAGAATGGTGAGCTCTGCCACCCG
-----+-----+-----+-----+-----+

15430

15450

15470

ACCCCTCCTTGCCGTTTGAATCCCGCAGGCCAGTGTCCCCCTCACTGCCTGGTGCACTGC
-----+-----+-----+-----+-----+

15490

15510

15530

CCGTAGGTGTGCCAGAGCAGGCGCTGCAGGAAGAATGCCCTTCCAGGAGCTTACGCGCTGC
-----+-----+-----+-----+-----+

15550

15570

15590

CTGACTGCCTGCCACAGCCACGGGGTGAGAGCCCGAGGAGTGGGGGTGACCTTGGGGTTC
-----+-----+-----+-----+-----+

15610

15630

15650

CTAATCCTACGTGACCCCTCCTCTTCTCTTCTCTGCAGGTTTGAATAGCAACCATAACTG
-----+-----+-----+-----+-----+

15670

15690

15710

CCACTGTGCTCCAGGCTGGGCTCCACCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCAT
-----+-----+-----+-----+-----+

15730

15750

15770

GGACAGTGGCCCTGTGTCAGGCTGAAAGTATGCCAGTGGGGGGCATGTGGGCAGGAGCTGG
-----+-----+-----+-----+-----+

15790

15810

15830

GGTGGTGCACCTGCTCAGGACTCAGCGCCCCTTCCCCCAATCCCCGCAGACCATGACACC
-----+-----+-----+-----+-----+

15850

15870

15890

TTCTGCTGGCCATGCTCCTCAGCGTCCTGCTGCCCTCTGCTCCCAGGGGCGCGCTGGCC
-----+-----+-----+-----+-----+

15910

15930

15950

TGGTGTGTGCTACCGACTCCAGGAGCCCATCTGCAGCGATGCACTGGGGCTGCAGAAGG
-----+-----+-----+-----+-----+

FIG. 29

122/157

```
15970      15990      16010
GACCCTGCCTGCAGTGGGTAGGCTCCGAGCGCCTGCTTCCTGAGCCTACTCCTGCGGTTCT
-----+-----+-----+-----+-----+-----+-----+
16030      16050      16070
CCCTCCTCAGAGCTCTGCTGGGGCTGTGGGAGCTGGGGCAGGCCCTCAGCCTTGCCCCCA
-----+-----+-----+-----+-----+-----+-----+
16090      16110      16130
GGTGCGAGAGAGCAGCCCCAGAGGCCATGGAAAGAAGTAGCTTTGAACAGGAGGTTCAGT
-----+-----+-----+-----+-----+-----+-----+
16150      16170      16190
GGCCTCCCAGTCAAGCGAGGGGGTGGATCCCTGCCCCACCACCAGCACCGCAAGGCATGG
-----+-----+-----+-----+-----+-----+-----+
16210      16230      16250
CCCTCTACCTCCCAGTACAGCTCCTCTTGTCCTACTCTCTGCTTCTCCCACCAGCTGGCT
-----+-----+-----+-----+-----+-----+-----+
16270      16290      16310
GCCTCACCTTTGACTTCGCCCTGTTTTTCCCTGGGCTCAGATTGCAGTCCCTGTACCATGC
-----+-----+-----+-----+-----+-----+-----+
16330      16350      16370
TGCCCCCGGAGGCCTGTCCAGCCTCTGTCTCACCAGTTTTGGGCCCTTTGCCACTTCCTC
-----+-----+-----+-----+-----+-----+-----+
16390      16410      16430
TGCACAAATCACCTCTGTCACCCCTTGAAGTTCCCAAATGCTGGGGCCCAGCACATCTTT
-----+-----+-----+-----+-----+-----+-----+
16450      16470      16490
TCACTCCATACCACTGGTCAGCTGGGTGCTGGCTGCCCTGTGCCAGGGGCCCTGCCTTA
-----+-----+-----+-----+-----+-----+-----+
16510      16530      16550
```

FIG. 29

124/157

```
17110      17130      17150
AGCCTGGAGGAAATGTTAGGACTAGAGAGAGGAGGTGCCAGCCGAGGGCATGAGGCTCAC
-----+-----+-----+-----+-----+-----+-----+
17170      17190      17210
TTGGAGCCTGGATCCCAAGGCTCCCTGAAGAGGGAGCAGGAAGGGAGCTGAGAGGGTGA
-----+-----+-----+-----+-----+-----+-----+
17230      17250      17270
CTTGGAGCAGATGGGTGCCCCAAGAACTCAGTAAACGCAGAACTCCCTGGGCTGGACAC
-----+-----+-----+-----+-----+-----+-----+
17290      17310      17330
CATGCTGCGGGGAGGCAATAACCCACTCAGGATCACTGTGCCAACCTCCTGGACTCTTAT
-----+-----+-----+-----+-----+-----+-----+
17350      17370      17390
CACGTTGCTCAGCCCCAAAGATGGCCACACAGGGACCACCCCCTGGCGGGCGTTACCC
-----+-----+-----+-----+-----+-----+-----+
17410      17430      17450
CATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCCCTGGGTGAGTGAGGCACCAG
-----+-----+-----+-----+-----+-----+-----+
17470      17490      17510
GGGGAGGTGGAGAGGGAAGGGAGAAGGGAAGGGCTCATGCCTCCTGCCTCCTTCCAGATG
-----+-----+-----+-----+-----+-----+-----+
17530      17550      17570
GGCAGCACCCAGTCACCTTGAGTCCCTATGCCCCCTCCCCAGCCCCAGGGTCTCCTGCTG
-----+-----+-----+-----+-----+-----+-----+
17590      17610      17630
ACCATATTACAAACATTTACCCTCCACCATTTCTCCAGACCTGAGAACTCTCATGAGC
-----+-----+-----+-----+-----+-----+-----+
17650      17670      17690
```

FIG. 29

CCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTCGCCTGACCCCCAAGGTAGGCAGG
- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +

17710 17730 17750
GACCTGGATTCAAAGCCTCCCCCTCTCATCGCCACCCCTCCACCTCTCCACCCCTCAG
- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +

17770 17790 17810
TTTGCTGCCCTAATCAGGTTTCTGGGCTCAGGTTATTATGGAAATGAGTTTATGACCT

17830 17850 17870

CTTGGTTATCATGGAGACCCAGGATGCTGGAAGCCCCCTGGGCTGGGGAGGGAGAAGCTGTG

17890 17910 17930

GCTTTTCCTGGATCACTGGTCCCTACTGAGTGAGGATGGGCTCTCTGCCACACAGCTTGC

- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +

17950 17970 17990
AGCCTGGGGCCCCAGTCTTAGGGGACAACATATCCTCCTCATTTCTCAGCAGATCAAGTCT

18010 18030 18050
CAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGCTCCTAAAATGAACAGATTAAAGAC

18070 18090 18110
AGGTGGCCACTGACAGCCACTCCAGGAAC TTGAACTGCAGGGGCAGAGCCAGTGAATCAC

CGGACCTCCAGCACCTGCAGGCAGCTTGGAA GTTCTCTCCCGAGTGGAGCTTGACCCA

18190 18210 18230
CCCACTCCAGGAACCCAGAGCCACATTAGAAGTTTCCTGAGGGCTGGAGAACACACTGCTGGG
-----+-----+-----+

FIG. 29

126/157

```
18250      18270      18290
CACACTCTCCAGCTCAATAAACCATCAGTCCCAGAAAGCAAAGGTCACACAGCCCCCTGACC
-----+-----+-----+-----+-----+-----+-----+
18310      18330      18350
TCCCTCACCAGTGGAGGCTGGGTAGTGTGGCCATCCCAAAAGGGCTCTGTCTGGGAGT
-----+-----+-----+-----+-----+-----+-----+
18370      18390      18410
CTGGTGTGTCTCTACATGCAATTTCACGGACCCAGCTCTGTGGAGGGCATGACTGTCTG
-----+-----+-----+-----+-----+-----+-----+
18430      18450      18470
GCCAGAAGCTAGTGGTCCTGGGGCCCTATGGTTCGACTGAGTCCACACTCCCCTGAGCC
-----+-----+-----+-----+-----+-----+-----+
18490      18510      18530
TGGCTGGCCTCTGCAAAACAAACATAATTTGGGGACCTTCTTCTCTGTTTCTTCCCACCC
-----+-----+-----+-----+-----+-----+-----+
18550      18570      18590
TGTCTTCTCCCTAGGTGGTTCCTGAGCCCCCACCCTCAATCCCACTGCTACACCTGAGG
-----+-----+-----+-----+-----+-----+-----+
18610      18630      18650
TTCTGGAGCTCAGAATCTGACAGCCTCTCCCCATTCTGTGTGTGTGTCGGGGGACAGAGG
-----+-----+-----+-----+-----+-----+-----+
18670      18690      18710
GAACCATTTAAGAAAAGATACCAAAGTAGAAGTCAAAGAAAGACATGTTGGCTATAGGC
-----+-----+-----+-----+-----+-----+-----+
18730      18750      18770
GTGGTGGCTCATGCCTATAATCCAGCACTTTGGGAAGCAGGGGTAGGAGGATCACCAGA
-----+-----+-----+-----+-----+-----+-----+
18790      18810      18830
```

FIG. 29

128/157

```
-----+-----+-----+-----+-----+-----+
      19390              19410              19430
CAAGCCCAAGGGGCCAGGTGAGAGTGCTCTCTAAAGCCCCCTCCTGCCTCAGTTGTAGTT
-----+-----+-----+-----+-----+-----+

      19450              19470              19490
GCAAAATTTTAATTTATGAAGGTGACTGATGACACAGAGGCCAATGCTGTTGAAATAAGT
-----+-----+-----+-----+-----+-----+

      19510              19530              19550
TATTACTCACAGTTTCCCAACCATGCAGGGCCACAGTGGGGAGGCACTAGGTTTGGTCCAG
-----+-----+-----+-----+-----+-----+

      19570              19590              19610
GGACAGAATCAGGAGCGAGTGGAAGGCACAGGCCACAGCCCACAGTGCCGTTTCCACTGG
-----+-----+-----+-----+-----+-----+

      19630              19650              19670
GGAGGCAAGGCAGGCCAGGGGAAGAGGGTAGGATTGGCATTTTGAATCATTCTGGTGGGG
-----+-----+-----+-----+-----+-----+

      19690              19710              19730
TTTGGGGCGTGGGGTTGGGCTCTAATTGTCTGGGTAGGTGCCTGGCCCTGAGCTGGTTTA
-----+-----+-----+-----+-----+-----+

      19750              19770              19790
GGGCAGGGGAAATACTTGGTTTCGTATGTGAGAGTTCCTTGAAGGGGGTGGTTGGTGTATG
-----+-----+-----+-----+-----+-----+

      19810              19830              19850
GACTCAAGACTGGTGGTTTGCATATGAAAGGCATGAGTTGTTTCTGATCTCCAGGAATC
-----+-----+-----+-----+-----+-----+

      19870              19890              19910
AAGCAGTTTCTCTCCAGCCAACAAGCCCCCACCCTGAGATGTTAAACCATCATAAAAATAG
-----+-----+-----+-----+-----+-----+
```

FIG. 29

129/157

| | | |
|---|-------|-------|
| 19930 | 19950 | 19970 |
| AGAATCTAAGGCCAGGCATGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAG | | |
| -----+-----+-----+-----+-----+-----+-----+ | | |
| 19990 | 20010 | 20030 |
| GCGGGAGGATCATTGAGGTCAGAAGTTCGAGACCAGCCTGGCCAAATGTGGTGAAACCCC | | |
| -----+-----+-----+-----+-----+-----+-----+ | | |
| 20050 | 20070 | 20090 |
| ATCTCTACTAAAAATACAAAAATTAGCCCGGTGTGGTGGCACGTGCCTATAATCCCAGCT | | |
| -----+-----+-----+-----+-----+-----+-----+ | | |
| 20110 | 20130 | 20150 |
| ACTCGGGAGGCTGCGGCAGGAGAAATTGTTTGAACATGGGAGGTGGAGGTTGCAGTGAGCT | | |
| -----+-----+-----+-----+-----+-----+-----+ | | |
| 20170 | 20190 | 20210 |
| GAGATCGTGCCACTGCACTCCAGCCTGGGCAACAAGAGCAAGACTCCGTCTCAAAAAAA | | |
| -----+-----+-----+-----+-----+-----+-----+ | | |
| 20230 | 20250 | 20270 |
| AAAAAAAAAAAAAGAGAGACTCTAAAAATACACGTTAATATACCTCCCCCGCTCTTACCCT | | |
| -----+-----+-----+-----+-----+-----+-----+ | | |
| 20290 | 20310 | 20330 |
| TCAGGAGGGGGTGTCTAGACCCCGCGGACTCCAGCTACAAGGGACCCTGGGGAGGCCAA | | |
| -----+-----+-----+-----+-----+-----+-----+ | | |
| 20350 | 20370 | 20390 |
| CTCTGCCCTCTTGGCTAATCCCCAAGACTGCCCAGCACCCCTCCACCCCTTCTCCATTCT | | |
| -----+-----+-----+-----+-----+-----+-----+ | | |
| 20410 | 20430 | 20450 |
| AGTGGCGAACCTGGGGAGGCCACGTGGGAAGGAAAGAGGGCTCTAAGAGGGGAGGCCCC | | |
| -----+-----+-----+-----+-----+-----+-----+ | | |
| 20470 | 20490 | 20510 |
| AGACTGGGGGAGAGGCCTGTCTGGAGCCCAGGATCACCTGGCTGTGCTGCAGAACTGGAG | | |

FIG. 29

130/157

```
-----+-----+-----+-----+-----+
      20530              20550              20570
AAGAGAAGCTCAGCAGAAAGGAGCTGGCATGGGGCCAACAGCAGAAAAGCAGGAGGCACG
-----+-----+-----+-----+-----+

      20590              20610              20630
CAGAAAGTGA CTGGGAAGCAGGAGGGTAGGCATGGACCCTGAGGCTGAGCAGGAGGTACTG
-----+-----+-----+-----+-----+

      20650              20670              20690
AGGGGCAGAGTGGACGCTGAGCTGGGGGTAGCGAGCGAGCCAGCTCAGCTGTGACGCCCC
-----+-----+-----+-----+-----+

      20710              20730              20750
TCTGTTTGGCCACCCA ACTACCAGCTACTTGGGCTGCCCCGGGAGGAACTGGGCTTCCTC
-----+-----+-----+-----+-----+

      20770              20790              20810
TGACATTCTGTGGCCTGCGGCCATCTGTACACCTTCTTCTCTCTGCCCCCTCCCTTGA
-----+-----+-----+-----+-----+

      20830              20850              20870
CTTGTTGGCACCCACAGACAGGTGGGAGAGTGTACCTGCCCTGTGTGGTCAGAGCTTGGTT
-----+-----+-----+-----+-----+

      20890              20910              20930
TTGAGTTTCCTTCCCTCACCCCTCTTTCCCTCCCACACGCCAAAAACACAAGAGGATGTGTC
-----+-----+-----+-----+-----+

      20950              20970              20990
AGAGGCCCTGTGAAC CAGAGCAACTCCATCCTGAATAGGGGCTGAGCAAAATAAGGCTGAG
-----+-----+-----+-----+-----+

      21010              21030              21050
ACCTACTGGGCTGCGTTTCCAGACAGTTACAGCATTCTGCGTCACAGGATGAGATAGGAG
-----+-----+-----+-----+-----+
```

FIG. 29

131/157

```
      21070      21090      21110
ATACAGGTCATAAAGACCTTGGCTGATAAAATAGTTTGCAGTAGGCCAGGCGCGGTAGCTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

      21130      21150      21170
ACGCCTGTAAATCCCAGCACTTTGGGAGGCTGAGGTGGGCGGATCACCTGAGGTCAGAAGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

      21190      21210      21230
TCGAGACCAGCCTGGCCAACAAGGTGAAACCTCATCTCTACTAAAAATACAAAAACTAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

      21250      21270      21290
CAGGCATGGTGGTGTGTGCCTGTAAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

      21310      21330      21350
CTTGAACCCAGGAGGTGGAGGTTGCAGTGAGCTGAGATCATGTCACTGCCTCCAGCCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

      21370      21390      21410
GAGCGAGACTCCGCTCTCAAAAACCAAAACCAACCAAAAAAATCAGCTTGCAATATAGAAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

      21430      21450      21470
TGGCTAAAACCCACCCAAACCAAGATGGTGATGAGAGTGACCTCTGGTCGGTCCCCACTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

      21490      21510      21530
CTACACTCCCACCAGCACCATGACAGGTTACAGATGCCATGGCAGTATCAGGAAGTTACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

      21550      21570      21590
ATATATGGTCTAAAAAGGGGAGACATGAACAATCCACCCCTGTTTAGCAGATCATCCAGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

      21610      21630      21650
AACCAACCATAAAAAATGGGCAACCAGCAGCCCTCAGGGCTGCGCTCTCTATGGAGTAGCCA
```

FIG. 29

132/157

```
-----+-----+-----+-----+-----+
      21670              21690              21710
TTCTTTTATCTTTTACTTTCTTAATAAATGTGCTTTCACTTTATGGACTCGTCTCAAAT
-----+-----+-----+-----+-----+

      21730              21750              21770
TCTTTCTTGACGAGATCCAAGAACCCTCTCCTGGGGTCTGAATCTGGACCCCTTTCCGG
-----+-----+-----+-----+-----+

      21790              21810              21830
TAACAGATGTCGTAGAGTGAAGCACAACTGCAGGGGCATCTGGTTTACATTTTGCT
-----+-----+-----+-----+-----+

      21850              21870              21890
TCAGCGCCCATGGTTAGCACAGCGGAAAGCACATCAGCTCTTCTGATTCATTAAAAAA
-----+-----+-----+-----+-----+

      21910              21930              21950
TTAGGAAATGGACCACCACAAACCACAGACAGATGTACTGAGACAGGATAGGTAGTCAAG
-----+-----+-----+-----+-----+

      21970              21990              22010
AAAGTGACCATGTTCTAGGCGCGCAGCAGCAACTGTGGTGACCGTACAGTCAACAAGCCT
-----+-----+-----+-----+-----+

      22030              22050              22070
CAGCACTGGCATTGCAATTGAGCTCATTTCAAGCAAAGCTATCTTCAGCAGGGACTTCTCC
-----+-----+-----+-----+-----+

      22090              22110              22130
CTCTAGGCAGCAAGCGCATTTTATTTTACCTGTCCTCAAAGTATGCTTTGCTCCTTAT
-----+-----+-----+-----+-----+

      22150              22170              22190
AACAGTAAGGAACACACCCCTGTGTGGAGATTTAAGATGCTAATGAGGCCAAGCGCAGTT
-----+-----+-----+-----+-----+
```

FIG. 29

133/157

```
22210      22230      22250
GCTCACGTCTGTAATTCCAGCACTTTGGGAGGCAGAGGTGGGCGGGTCACTTGAGGTTAG
-----+-----+-----+-----+-----+-----+-----+

22270      22290      22310
AAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTTGTCCTACTAAAAATACAAAAAT
-----+-----+-----+-----+-----+-----+-----+

22330      22350      22370
TAGCCGGGCATGGTGGCGGCGCTCTGTAATCCAGCTACCTGGGAGGCTGAGGCAGAAGA
-----+-----+-----+-----+-----+-----+-----+

22390      22410      22430
ATCGCTTGAACCTGGGAGGCGGAGGTTGCAGTGAGCCAAGATCGTGCCACTGCACTCCAG
-----+-----+-----+-----+-----+-----+-----+

22450      22470      22490
CCTGAGGGGAGAGAGAGAGCAAGACATCGTTTTTGTGTTGTTGTTGTTGTTGTTGTT
-----+-----+-----+-----+-----+-----+-----+

22510      22530      22550
TTTTAAAAAAGTCAAGACAAATCATAGTGGGGGCTTTTCTGGTCACTTTTTAAATCTT
-----+-----+-----+-----+-----+-----+-----+

22570      22590      22610
AGTGTGAGACTTTATTTGAGACAGGCCTTACTCTGTTGCCAGGTTGGATGAGATTTT
-----+-----+-----+-----+-----+-----+-----+

22630      22650      22670
TAACCTCAATATTTACTTATAGAATAACTTTTTGTTAGTCAAAACAATGCTGTGCTCTCA
-----+-----+-----+-----+-----+-----+-----+

22690      22710      22730
TTCTGATCAGAATAAAACATCAGACAACTCAAGAGAAACATTCTGCAAAATAACTGGCCA
-----+-----+-----+-----+-----+-----+-----+

22750      22770      22790
GGATTCTTCAAAAGTGTCAAGGGTAAAGATAAGGAAAGATGAAGGAACCTCCAGATTGAG
```

FIG. 29

134/157

```
-----+-----+-----+-----+-----+
      22810              22830              22850
GAGAATAAGGAGACAACACTGTGATGTGGGATCCTAGAAATGGATCTTGGAAACAGAAAAGGA
-----+-----+-----+-----+-----+

      22870              22890              22910
CATTAGTGGAAAAATGAGAAATGCAAAACAGTCTACAGTTTCGTTAACAGGATTGTACCA
-----+-----+-----+-----+-----+

      22930              22950              22970
AGGTTAGTTTCCTAGCTGTAATGATTGGACTATGATTAAAGTAAGATGGACCATCAGGGGA
-----+-----+-----+-----+-----+

      22990              23010              23030
AGCTGGGTGAAGGGTGTAAGGAAAATGCTTACATTTTCCAACTTTTCTGCAAGTCTAAAA
-----+-----+-----+-----+-----+

      23050              23070              23090
TTAGTCAACAATAAGAAGTTTAAAAATAGGCCAGGCATGGTGGCTCACACCTGTCATCCTA
-----+-----+-----+-----+-----+

      23110              23130              23150
GCACTTTGAGAGGCCCGAGGTGGGAGGATGGTTTGAGCCCAGGAGTTCAAGACCAGCCTGC
-----+-----+-----+-----+-----+

      23170              23190              23210
GCAATAGAGCGAGACCCCAACTCTATTCAAAAAAATTTTTTAAGTTTAAAAATAGAATTA
-----+-----+-----+-----+-----+

      23230              23250              23270
TATAAAAAAGGAAAAGAAAATGCTGTTTCATAGCGTTCCTAGTTTAGCATGGGAGAGAC
-----+-----+-----+-----+-----+

      23290              23310              23330
CAGGTCTCCCTGGGTGGTTGCTCTGTGTGTTGCTGGGTGTGCGTGGCAGGGCTAGTGTGTT
-----+-----+-----+-----+-----+
```

FIG. 29

135/157

| | | |
|---|-------|-------|
| 23350 | 23370 | 23390 |
| GGGGTCCGTCTAGGCACATTCAGGCGCCGAATCCCGTGGCTCCCAGGTTTACCTGACGGT | | |
| -----+-----+-----+-----+-----+-----+ | | |
| 23410 | 23430 | 23450 |
| GCAGCCTGGGGTGGAGACTTAATGAGGGCGGGGAGTTGCTGCAGCAAAGGCTCCTCCCAG | | |
| -----+-----+-----+-----+-----+-----+ | | |
| 23470 | 23490 | 23510 |
| GGGTATCAGCGCAGACAGCTGGGTTTTCACTGTGCTCCTGCTCCAGAGGCAC TAGGAAGG | | |
| -----+-----+-----+-----+-----+-----+ | | |
| 23530 | 23550 | 23570 |
| GGGCGCCTATCAGACTAGGACTCTGCCAGCCATCCTTCTCTGTTGAAGGTCCAGC | | |
| -----+-----+-----+-----+-----+-----+ | | |

FIG. 29

136/157

```

      10              30              50
CAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCCCGTTGCTGCTGCTGCTACTAC
-----+-----+-----+-----+-----+
MetGlyTrpArgProArgArgAlaArgGlyThrProLeuLeuLeuLeuLeuLeuL

      70              90              110
TGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGACATATCCCTGGGCAGC
-----+-----+-----+-----+-----+
euLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGlyHisIleProGlyGlnP

      130             150             170
CAGTCACCCCGCACTGGGTCTGGATGGACAACCCCTGGCGCACCGTCAGCCTGGAGGAGC
-----+-----+-----+-----+-----+
roValThrProHisTrpValLeuAspGlyGlnProTrpArgThrValSerLeuGluGluP

      190             210             230
CGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAAGGCCAGGAGCTCCTGC
-----+-----+-----+-----+-----+
roValSerLysProAspMetGlyLeuValAlaLeuGluAlaGluGlyGlnGluLeuLeuL

      250             270             290
TTGAGCTGGAGAAGAACCACAGGCTGCTGGCCCCAGGATACATAGAAACCCACTACGGCC
-----+-----+-----+-----+-----+
euGluLeuGluLysAsnHisArgLeuLeuAlaProGlyTyrIleGluThrHisTyrGlyP

      310             330             350
CAGATGGGCAGCCAGTGGTGTGGCCCCCAACCACACGGATCATTGCCACTACCAAGGGC
-----+-----+-----+-----+-----+
roAspGlyGlnProValValLeuAlaProAsnHisThrAspHisCysHisTyrGlnGlyA

      370             390             410
GAGTAAGGGGCTTCCCCGACTCCTGGGTAGTCCTCTGCACCTGCTCTGGGATGAGTGGCC
-----+-----+-----+-----+-----+
rgValArgGlyPheProAspSerTrpValValLeuCysThrCysSerGlyMetSerGlyL

      430             450             470
TGATCACCCCTCAGCAGGAATGCCAGCTATTATCTGCGTCCTTGGCCACCCCGGGGCTCCA
-----+-----+-----+-----+-----+
euIleThrLeuSerArgAsnAlaSerTyrTyrLeuArgProTrpProProArgGlySerL

```

FIG. 30

137/157

```

      490              510              530
AGGACTTCTCAACCCACGAGATCTTTTCGGATGGAGCAGCTGCTCACCTGGAAAGGAACCT
-----+-----+-----+
ysAspPheSerThrHisGluIlePheArgMetGluGlnLeuLeuThrTrpLysGlyThrC

      550              570              590
GTGGCCACAGGGATCCTGGGAACAAAGCGGGCATGACCAGCCTTCTGGTGGTCCCCAGA
-----+-----+-----+
ysGlyHisArgAspProGlyAsnLysAlaGlyMetThrSerLeuProGlyGlyProGlnS

      610              630              650
GCAGGGGCAGGCGAGAAAGCGCGCAGGACCCGGAAGTACCTGGAAGTGTACATTGTGGCAG
-----+-----+-----+
erArgGlyArgArgGluAlaArgArgThrArgLysTyrLeuGluLeuTyrIleValAlaA

      670              690              710
ACCACACCCCTGTTCTTGACTCGGCACCGAAACTTGAACCAACCAACAGCGTCTCCTGG
-----+-----+-----+
spHisThrLeuPheLeuThrArgHisArgAsnLeuAsnHisThrLysGlnArgLeuLeuG

      730              750              770
AAGTCGCCAACTACGTGGACCAGCTTCTCAGGACTCTGGACATTACAGGTGGCGCTGACCG
-----+-----+-----+
luValAlaAsnTyrValAspGlnLeuLeuArgThrLeuAspIleGlnValAlaLeuThrG

      790              810              830
GCCTGGAGGTGTGGACCGAGCGGGACCGCAGCCGCGTCACGCAGGACGCCAACGCCACGC
-----+-----+-----+
lyLeuGluValTrpThrGluArgAspArgSerArgValThrGlnAspAlaAsnAlaThrL

      850              870              890
TCTGGGCCTTCTGCAGTGGCGCCGGGGCTGTGGGCGCAGCGGCCCAACGACTCCGCGC
-----+-----+-----+
euTrpAlaPheLeuGlnTrpArgArgGlyLeuTrpAlaGlnArgProHisAspSerAlag

      910              930              950
AGCTGCTCACGGGCCGCGCCTTCCAGGGCGCCACAGTGGGCCTGGCGCCCGTCGAGGGCA
-----+-----+-----+
lnLeuLeuThrGlyArgAlaPheGlnGlyAlaThrValGlyLeuAlaProValGluGlyM

```

FIG. 30

138/157

```

          970                      990                      1010
TGTGCCGCGCCGAGAGCTCGGGAGGCGGTGAGCACGGACCACTCGGAGCTCCCCATCGGCG
-----+-----+-----+-----+-----+-----+-----+
etCysArgAlaGluSerSerGlyGlyValSerThrAspHisSerGluLeuProIleGlyA

          1030                      1050                      1070
CCGCAGCCACCATTGGCCCATGAGATCGGCCACAGCCTCGGCCTCAGCCACGACCCCGACG
-----+-----+-----+-----+-----+-----+-----+
1aAlaAlaThrMetAlaHisGluIleGlyHisSerLeuGlyLeuSerHisAspProAspG

          1090                      1110                      1130
GCTGCTGCGTGGAGGCTGCGGCCGAGTCCGGAGGCTGCGTCATGGCTGCGGCCACCGGGC
-----+-----+-----+-----+-----+-----+-----+
lyCysCysValGluAlaAlaAlaGluSerGlyGlyCysValMetAlaAlaAlaThrGlyH

          1150                      1170                      1190
ACCCGTTTCGCGCGTGTTCAGCGCCTGCAGCCGCCGCCAGCTGCGCGCGCTTCTTCCGCA
-----+-----+-----+-----+-----+-----+-----+
isProPheProArgValPheSerAlaCysSerArgArgGlnLeuArgAlaPhePheArgL

          1210                      1230                      1250
AGGGGGGCGGCGCTTGCTCTCCAAATGCCCCGACCCCGACTCCCGGTGCCGCCGGCGC
-----+-----+-----+-----+-----+-----+-----+
ysGlyGlyGlyAlaCysLeuSerAsnAlaProAspProGlyLeuProValProProAlaL

          1270                      1290                      1310
TCTGCGGGAACGGCTTCGTGGAAGCGGGCGAGGAGTGTGACTGCGGCCCTGGCCAGGAGT
-----+-----+-----+-----+-----+-----+-----+
euCysGlyAsnGlyPheValGluAlaGlyGluGluCysAspCysGlyProGlyGlnGluC

          1330                      1350                      1370
GCCGCGACCTTGCTGCTTGTGCTCACAAC TGCTGCTGCGCCCGGGGGGCCAGTGCGCCC
-----+-----+-----+-----+-----+-----+-----+
ysArgAspLeuCysCysPheAlaHisAsnCysSerLeuArgProGlyAlaGlnCysAlaH

          1390                      1410                      1430
ACGGGGACTGCTGCGTGCGCTGCTGCTGAAGCCGGCTGGAGCGCTGTGCCCGCCAGGCCA
-----+-----+-----+-----+-----+-----+-----+
isGlyAspCysCysValArgCysLeuLeuLysProAlaGlyAlaLeuCysArgGlnAlaM

          1450                      1470                      1490

```

FIG. 30

139/157

```

TGGGTGACTGTGACCTCCCTGAGTTTTCACGGGCACCTCCTCCCACTGTCCCCCAGACG
-----+-----+-----+-----+-----+-----+
etGlyAspCysAspLeuProGluPheCysThrGlyThrSerSerHisCysProProAspV

1510              1530              1550
TTTACCTACTGGACGGCTCACCCCTGTGCCAGGGGCAGTGGCTACTGTGGGATGGCGCAT
-----+-----+-----+-----+-----+-----+
alTyrLeuLeuAspGlySerProCysAlaArgGlySerGlyTyrCysTrpAspGlyAlaC

1570              1590              1610
GTCCACGCTGGAGCAGCAGTGCCAGCAGCTCTGGGGGCCTGGCTCCACCCAGCTCCCG
-----+-----+-----+-----+-----+-----+
ysProThrLeuGluGlnGlnCysGlnGlnLeuTrpGlyProGlySerHisProAlaProG

1630              1650              1670
AGGCCTGTTTCCAGGTGGTGAACCTCTGCGGGAGATGCTCATGGAAACTGCGGCCAGGACA
-----+-----+-----+-----+-----+-----+
luAlaCysPheGlnValValAsnSerAlaGlyAspAlaHisGlyAsnCysGlyGlnAspS

1690              1710              1730
GCGAGGGCCACTTCTCGCCCTGTGCAGGGAGGGATGCCCTGTGTGGGAAGCTGCAGTGCC
-----+-----+-----+-----+-----+-----+
erGluGlyHisPheLeuProCysAlaGlyArgAspAlaLeuCysGlyLysLeuGlnCysG

1750              1770              1790
AGGGTGGAAGCCAGCCTGCTCGCACCGCACATGGTGCCAGTGGACTTACCGTTCACC
-----+-----+-----+-----+-----+-----+
lnGlyGlyLysProSerLeuLeuAlaProHisMetValProValAspSerThrValHisL

1810              1830              1850
TAGATGGCCAGGAAGTGACTTGTGCGGGAGCCTTGGCACTCCCCAGTGCCAGCTGGACC
-----+-----+-----+-----+-----+-----+
euAspGlyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLeuAspL

1870              1890              1910
TGCTTGGCCTGGGCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAATGGTGTGCCAGA
-----+-----+-----+-----+-----+-----+
euLeuGlyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCysGlnS

1930              1950              1970

```

FIG. 30

140/157

```

GCAGCGCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTGCCACA
-----+-----+-----+-----+-----+
erArgArgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCysHisS

      1990              2010              2030
GCCACGGGGTTTGC AATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCCACCTT
-----+-----+-----+-----+
erHisGlyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaProProP

      2050              2070              2090
TCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGAAAACC
-----+-----+-----+-----+
heCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGluAsnH

      2110              2130              2150
ATGACACCTTCCTGCTGGCCATGCTCCTCAGCGTCCTGCTGCCTCTGCTCCCAGGGGCCG
-----+-----+-----+-----+
isAspThrPheLeuLeuAlaMetLeuLeuSerValLeuLeuProLeuLeuProGlyAlaG

      2170              2190              2210
GCCTGGCCTGGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTGGGGCT
-----+-----+-----+-----+
lyLeuAlaTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTrpGlyC

      2230              2250              2270
GCAGAAGGGACCTGCGTGCAGTGGCCCCAAAGATGGCCCCACACAGGGACCACCCCTGG
-----+-----+-----+-----+
ysArgArgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisProLeuG

      2290              2310              2330
GCGGCGTTTACCCCCATGGAGTTGGGCCCCACAGCCACTGGACAGGCCTGGCCCCCTGGACC
-----+-----+-----+-----+
lyGlyValHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLeuAspP

      2350              2370              2390
CTGAGAACTCTCATGAGCCCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTCGCCTG
-----+-----+-----+-----+
roGluAsnSerHisGluProSerSerHisProGluLysProLeuProAlaValSerProA

      2410              2430              2450
ACCCCCAAGATCAAGTCCAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGCTCCTAAAAA

```

FIG. 30

141/157

-----+-----+-----+-----+-----+
spProGlnAspGlnValGlnMetProArgSerCysLeuTrpEnd

2470 2490 2510
TGAACAGATTTAAAGACAGGTGGCCACTGACAGCCACTCCAGGAACCTGAACTGCAGGGG
-----+-----+-----+-----+-----+

2530 2550 2570
CAGAGCCAGTGAATCACCGACCTCCAGCACCTGCAGGCAGCTTGGAAAGTTCTTCCTCCCG
-----+-----+-----+-----+-----+

2590 2610 2630
AGTGGAGCTTCGACCCACCCACTCCAGGAACCCAGAGCCACATTAGAAAGTTCCTGAGGGC
-----+-----+-----+-----+-----+

2650 2670 2690
TGGAGAACACTGCTGGGCACACTCTCCAGCTCAATAAACCATCAGTCCCAGAAGCAAAGG
-----+-----+-----+-----+-----+

2710 2730 2750
TCACACAGCCCCTGACCTCCCTCACCAGTGGAGGCTGGGTAGTGCTGGCCATCCCAAAAG
-----+-----+-----+-----+-----+

2770 2790 2810
GGCTCTGTCTGGGAGTCTGGTGTGTCTCCTACATGCAATTCCACGGACCCAGCTCTGT
-----+-----+-----+-----+-----+

2830 2850 2870
GGAGGGCATGACTGCTGGCCAGAAGCTAGTGGTCCTGGGGCCCTATGGTTCGACTGAGTC
-----+-----+-----+-----+-----+

2890 2910 2930
CACACTCCCCTGCAGCCTGGCTGGCCTCTGCAAACAAACATAAATTTGGGGACCTTCCTT
-----+-----+-----+-----+-----+

2950 2970 2990
CCTGTTTCTTCCCACCCTGTCTTCTCCCTAGGTGGTTCCTGAGCCCCCACCCCAATCC
-----+-----+-----+-----+-----+

FIG. 30

142/157

```

3010          3030          3050
CAGTGTCTACACCTGAGGTTCTGGAGCTCAGAATCTGACAGCCTCTCCCCATTCTGTGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+

3070          3090          3110
TGTGCGGGGGACAGAGGGGAACCATTTAAGAAAAGATACCAAAGTAGAAGTCAAAAGAAAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+

3130          3150          3170
ACATGTTGGCTATAGGCGTGGTGGCTCATGCCTATAATCCCAGCACTTTGGGAAGCCGGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+

3190          3210          3230
GTAGGAGGATCACCAGAGGCCAGCAGGTCCACACCAGCCTGGGCAACACAGCAAGACACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+

3250          3270          3290
GCATCTACAGAAAAATTTTAAAAATTAGCTGGGCGTGGTGGTGTGTACCTGTAGGCCTAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+

3310          3330          3350
TGCTCAGGAGGCTGAAGCAGGAGGATCACTTGAGCCTGAGTTCAACACTGCAGTGAGCTA
-----+-----+-----+-----+-----+-----+-----+-----+-----+

3370          3390          3410
TGGTGGCACCACCTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCTAAAATAAAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+

3430          3450          3470
TTTAAAAAGACATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+

3490
AAAAAAAAAAAAAAAAAAAAAAAAAAAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

FIG. 30

143/157

10 30 50
CGGGCACGGGTCGGCCGCAATCCAGCCTGGGGCGGAGCCGGAGTTGCGAGCCGCTGCCTAG

70 90 110
AGGCCGAGGAGCTCACAGCTATGGGGCTGGAGGCCCCGGAGAGCTCGGGGGACCCCGTTGC
MetGlyTrpArgProArgArgAlaArgGlyThrProLeuL

130 150 170
TGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGAC
euLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGlyH

190 210 230
ATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCCTGGATGGACAACCCCTGGCGCACCG
isIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpArgThrV

250 270 290
TCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGGCTGAAG
alSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluAlaGluG

310 330
GCCAGGAGCTCCTGCTTGAGCTGGAGAAGAACCACAGGC
lyGlnGluLeuLeuLeuGluLeuGluLysAsnHisArg

FIG. 31

144/157

10 30 50
CGGGCACGGGTTCGGCCGCAATCCAGCCTGGGCGGAGCCGGAGTTGCGAGCCGCTGCCTAG

70 90 110
AGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGACCCCGTTGC
MetGlyTrpArgProArgArgAlaArgGlyThrProLeuL

130 150 170
TGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGAC
euLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGlyH

190 210
ATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCCTGGATGGAC
isIleProGlyGlnProValThrProHisTrpValLeuAspGly

FIG. 32

145/157

10 30 50
GCCTAGAGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCC
MetGlyTrpArgProArgArgAlaArgGlyThrP

70 90 110
CGTTGCTGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTC
roLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuG

130 150 170
AAGGACATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCCTGGATGGACAACCCCTGGC
lnGlyHisIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpA

190 210 230
GCACCGTCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGCTGGTGGCCCTGGAGG
rgThrValSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluA

250 270 290
CTGAAGGCCAGGAGCTCCTGCTTGAGCTGGAGAACAACACAGGCTGCTGGCCCCAGGAT
laGluGlyGlnGluLeuLeuLeuGluLeuGluLysAsnHisArgLeuLeuAlaProGlyT

310 330 350
ACATAGAAACCCACTACGGCCAGATGGGCAGCCAGTGGTGCTGGCCCCAACACACGG
yrIleGluThrHisTyrGlyProAspGlyGlnProValValLeuAlaProAsnHisThrV

370 390 410
TGAGATGCTTCCATGGGCTCTGGGATGCACCGCCAGAGGATCATTGCCACTACCAAGGGC
alArgCysPheHisGlyLeuTrpAspAlaProProGluAspHisCysHisTyrGlnGlyA

430 450 470
GAGTAAGGGGCTTCCCCGACTCCTGGGTAGTCCTCTGCACCTGCTCTGGGATGAGTGGCC
rgValArgGlyPheProAspSerTrpValValLeuCysThrCysSerGlyMetSerGlyL

490 510 530
TGATCACCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGGCCACCCCGGGCTCCA
euIleThrLeuSerArgAsnAlaSerTyrTyrLeuArgProTrpProProArgGlySerL

550
AGGACTTCTCAACCCACGAGAT
ysAspPheSerThrHisGlu

FIG. 33

146/157

10 30 50
GAGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCCCGTTG
MetGlyTrpArgProArgArgAlaArgGlyThrProLeu

70 90 110
CTGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGTGCTTCAAGGA
LeuLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGly

130 150 170
CATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCCTGGATGGACAACCCCTGGCGCACC
HisIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpArgThr

190 210 230
GTCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAA
ValSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluAlaGlu

250 270 290
GGCCAGGAGCTCCTGCTTGAGCTGGAGAAGAACCATGGCCTGATCACCCCTCAGCAGGAAT
GlyGlnGluLeuLeuLeuGluLeuGluLysAsnHisGlyLeuIleThrLeuSerArgAsn

310 330 350
GCCAGCTATTATCTGCGTCCCTGGCCACCCCGGGGCTCCAAGGACTTCTCAACCCACGAG
AlaSerTyrTyrLeuArgProTrpProProArgGlySerLysAspPheSerThrHisGlu

AT

FIG. 34

147/157

10 30 50
GAGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCCCGTTG
MetGlyTrpArgProArgAlaArgGlyThrProLeu

70 90 110
CTGTGCTGTACTACTGTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGA
LeuLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGly

130 150 170
CATATCCCTGGGCAGCCAGTCACCCGCACTGGGTCCTGGATGGACAACCTGGCGCACCC
HisIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpArgThr

190 210 230
GTCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAA
ValSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluAlaGlu

250 270 290
GGCCAGGAGCTCCTGTCTTGAGCTGGAGAAGAACCACAGGCTGCTGGCCCCAGGATACATA
GlyGlnGluLeuLeuLeuGluLeuGluLysAsnHisArgLeuLeuAlaProGlyTyrIle

310 330 350
GAAACCCACTACGGCCCAGATGGGCAGCCAGTGGTGTCTGGCCCCCAACCACACGGATCAT
GluThrHisTyrGlyProAspGlyGlnProValValLeuAlaProAsnHisThrAspHis

370 390 410
TGCCACTACCAAGGGCGAGTAAGGGGCTTCCCCGACTCCTGGGTAGTCCTCTGCACCTGC
CysHisTyrGlnGlyArgValArgGlyPheProAspSerTrpValValLeuCysThrCys

430 450 470
TCTGGGATGAGTGGCCTGATCACCCCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGG
SerGlyMetSerGlyLeuIleThrLeuSerArgAsnAlaSerTyrTyrLeuArgProTrp

490 510
CCACCCCGGGGCTCCAAGGACTTCTCAACCCACGAGAT
ProProArgGlySerLysAspPheSerThrHisGlu

FIG. 35

148/157

10 30 50
 CTGGCCCCAGGATACATAGAAACCCACTACGGCCCAGATGGGCAGCCAGTGGTGCTGGCC
 LeuAlaProGlyTyrIleGluThrHisTyrGlyProAspGlyGlnProValValLeuAla

70 90 110
 CCCAACCACACGGATCATTGCCACTACCAAGGGCGAGTAAGGGGCTTCCCCGACTCCTGG
 ProAsnHisThrAspHisCysHisTyrGlnGlyArgValArgGlyPheProAspSerTrp

130 150 170
 GTAGTCCTCTGCACCTGCTCTGGGATGAGTGGCCTGATCACCCCTCAGCAGGAATGCCAGC
 ValValLeuCysThrCysSerGlyMetSerGlyLeuIleThrLeuSerArgAsnAlaSer

190 210 230
 TATTATCTGCGTCCCTGGCCACCCCGGGCTCCAAGGACTTCTCAACCCACGAGATCTTT
 TyrTyrLeuArgProTrpProProArgGlySerLysAspPheSerThrHisGluIlePhe

250 270 290
 CGGATGGAGCAGCTGCTCACCTGGAAAGGAACCTGTGGCCACAGGGATCCTGGGGAACAAA
 ArgMetGluGlnLeuLeuThrTrpLysGlyThrCysGlyHisArgAspProGlyAsnLys

310 330 350
 GCGGGCATGACCAGCCTTCCTGGTGGTCCCCAGAGCAGGGGCAGGCGAAAAGCGCGCAGG
 AlaGlyMetThrSerLeuProGlyGlyProGlnSerArgGlyArgArgLysAlaArgArg
 Glu

370 390 410
 ACCCGGAAGTACCTGGAACGTACATTGTGGCAGACCACACCCTGTCTTGGACTCGGCAC
 ThrArgLysTyrLeuGluLeuTyrIleValAlaAspHisThrLeuPheLeuThrArgHis

430 450 470
 CGAAACTTGAACCACACCAACAGCGTCTCCTGGAAGTCGCCAACTACGTGGACCAGCTT
 ArgAsnLeuAsnHisThrLysGlnArgLeuLeuGluValAlaAsnTyrValAspGlnLeu

490
 CTCAGGACTCTGGACATTTCAGGTGGC
 LeuArgThrLeuAspIleGlnVal

FIG. 36

149/157

10 30 50
CAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

70 90 110
GGGGCTGGCTCCCACCCAGCTCCCGAGGCCTGTTTCCAGGTGGTGAAGTCTGCGGGAGA
pGlyProGlySerHisProAlaProGluAlaCysPheGlnValValAsnSerAlaGlyAs

130 150 170
TGCTCATGGAACTGCGGCCAGGACAGCGAGGGCCACTTCCTGCCCTGTGCAGGGAGGGA
pAlaHisGlyAsnCysGlyGlnAspSerGluGlyHisPheLeuProCysAlaGlyArgAs

190 210 230
TGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTGGAAGCCCAGCCTGCTCGCACCGCACAT
pAlaLeuCysGlyLysLeuGlnCysGlnGlyGlyLysProSerLeuLeuAlaProHisMe

250 270 290
GGTGCCAGTGGAAGCTTACCGTTACCTAGATGGCCAGGAAGTGACTTGTGCGGGAGCCTT
tValProValAspSerThrValHisLeuAspGlyGlnGluValThrCysArgGlyAlaLe

310 330 350
GGCACTCCCCAGTGCCAGCTGGACCTGCTTGGCCTGGGCCTGGTAGAGCCAGGCACCCA
uAlaLeuProSerAlaGlnLeuAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGl

370 390 410
GTGTGGACCTAGAAATGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGC
nCysGlyProArgMetValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAl

430 450 470
TCCACCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGC
aProProPheCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAl

490
TGAAAACCATGACACCTTCCTGC
aGluAsnHisAspThrPheLeu

FIG. 37

150/157

```

      10              30              50
CAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

      70              90              110
GGGGCCTGGCTCCACCCAGCTCCCAGGCGCTGTTTCCAGGTGGTGAACCTCTGCGGGAGA
pGlyProGlySerHisProAlaProGluAlaCysPheGlnValValAsnSerAlaGlyAs

      130             150             170
TGCTCATGGAAC TGCGGCCAGGACAGCGAGGGCCACTTCCTGCCCTGTGCAGGGAGGGA
pAlaHisGlyAsnCysGlyGlnAspSerGluGlyHisPheLeuProCysAlaGlyArgAs

      190             210             230
TGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTGGAAAGCCAGCCTGCTGCACCGCACAT
pAlaLeuCysGlyLysLeuGlnCysGlnGlyGlyLysProSerLeuLeuAlaProHisMe

      250             270             290
GGTGCCAGCTGGACTCTACCGTTACCTAGATGGCCAGGAAGTGACTTGTCGGGGAGCCTT
tValProValAspSerThrValHisLeuAspGlyGlnGluValThrCysArgGlyAlaLe

      310             330             350
GGCAC TCCCAGT GCCAGCTGGAC TGT TGGCCTGGGCCTGGTAGAGCCAGGCACCCA
uAlaLeuProSerAlaGlnLeuAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGl

      370             390             410
GTGTGGAC TAGAATGGTGTGCCAGAGCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCT
nCysGlyProArgMetValCysGlnSerArgArgCysArgLysAsnAlaPheGlnGluLe

      430             450             470
TCAGCGCTGCCTGACTGCCTGCCACAGCCACGGGGTTTGCATAGCAACCATAACTGCCA
uGlnArgCysLeuThrAlaCysHisSerHisGlyValCysAsnSerAsnHisAsnCysHi

      490             510             530
CTGTGCTCCAGGCTGGGCTCCACCCTTCTGTGACAAGCCAGGCTTGTGTGCAGCATGGGA
scysAlaProGlyTrpAlaProProPheCysAspLysProGlyPheGlyGlySerMetAs

      550             570
CAGTGGCCCTGTGCAGGCTGAAAACCATGACACCTTCCTGC
pSerGlyProValGlnAlaGluAsnHisAspThrPheLeu

```

FIG. 38

151/157

10 30 50
CAGTGGCTACTGCTGGGATGGCGCATGTCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

70 90 110
GGGGCCTGATGGCCAGGAAGTGACTTGTGCGGGGAGCCTTGGCACTCCCCAGTGCCACAGT
pGlyProAspGlyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLe

130 150 170
GGACCTGCTTGGCCTGGGCCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAATGGTGTG
uAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCy

190 210 230
CCAGAGCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTG
sGlnSerArgArgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCy

250 270 290
CCACAGCCACGGGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCC
sHisSerHisGlyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaPr

310 330 350
ACCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGA
oProPheCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGl

370
AAACCATGACACCTTCCTGC
uAsnHisAspThrPheLeu

FIG. 39

152/157

10 30 50
GGCCTGCTGTTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTGGGGCTGCAG
AlaTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTrpGlyCysAr

70 90 110
AAGGGACCCCTGCGTGTCAGTGGCCCCAAAGATGGCCACACAGGGACCACCCCTGGGCGG
gArgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisProLeuGlyGl

130 150 170
CGTTCACCCCATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCTGGACCCTGA
yValHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLeuAspProGl

190 210 230
GAACTCTCATGAGCCCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTCGCCTGACCC
uAsnSerHisGluProSerSerHisProGluLysProLeuProAlaValSerProAspPr

250 270 290
CCAAGCAGATCAAGTCCAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGCTCCTAAAT
oGlnAlaAspGlnValGlnMetProArgSerCysLeuTrpEnd

310
GAACAGATTTAAAGACAGGTGGCC

FIG. 40

153/157

```

      10              30              50
CAGTGGCTACTGCTGGGATGGCGCATGTCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

      70              90              110
GGGCGCTGATGGCCAGGAAGTGACTTGTGGGGAGCCCTTGGCACTCCCCAGTGCCAGCT
pGlyProAspGlyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLe

     130              150              170
GGACCTGCTTGGCCTGGGCGCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAATGTTGTG
uAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCy

     190              210              230
CCAGAGCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTG
sGlnSerArgArgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCy

     250              270              290
CCACAGCCAGCGGGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCC
sHisSerHisGlyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaPr

     310              330              350
ACCCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGA
oProPheCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGl

     370              390              410
AAACCATGACACCTTCTGCTGGCCATGCTCCTCAGCGTCTCTGCTGCCTCTGCTCCAGG
uAsnHisAspThrPheLeuLeuAlaMetLeuLeuSerValLeuLeuProLeuLeuProGl

     430              450              470
GGCCGGCGCTGGCCTGGTGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTG
yAlaGlyLeuAlaTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTr

     490              510              530
GGGCTGCAGAAAGGGACCOCTGCGTGCAGTGGCCCCAAGATGGCCACACAGGGACCCACC
pGlyCysArgArgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisPr

     550              570              590
CCTGGGCGGCGTTACGCCATGGAGTTGGGGCCCCACAGCCACTGGACAGCCCTGGCCCTC
oLeuGlyGlyValHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLe

     610              630              650
GGACCTTGAGAAGTCTCATGAGCCAGCAGCCACCCCTGAGAAGCTCTGCCAGCAGTCTC
uAspProGluAanSerHisGluProSerSerHisProGluLysProLeuProAlaValse

     670              690              710
GCCTGACCCCCAAGCAGATCAAGTCCAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGC
rProAspProGlnAlaAspGlnValGlnMetProArgSerCysLeuTrpEnd

     730              750
TCCTAAATGAACAGATTAAAGACAGGTGGCC

```

FIG. 41

154/157

10 30 50
 CAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGAGCAGCAGTGGCCAGCAGCTCTG
 SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr
 70 90 110
 GGGGCTGGCTCCCACCCAGCTCCCAGGCCTGTTTCCAGGTGGTGAACCTCTGCGGGAGA
 pGlyProGlySerHisProAlaProGluAlaCysPheGlnValValAsnSerAlaGlyAs
 130 150 170
 TGCTCATGGAACTGCGGCCAGGACAGCGAGGGCCACTTCCTGCCCTGTGCAGGGAGGGA
 pAlaHisGlyAsnCysGlyGlnAspSerGluGlyHisPheLeuProCysAlaGlyArgAs
 190 210 230
 TGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTGGAAAGCCCAGCCTGCTCGCACCCGACAT
 pAlaLeuCysGlyLysLeuGlnCysGlnGlyGlyLysProSerLeuLeuAlaProHisMe
 250 270 290
 GGTGCCAGTGGACTCTACCGTTCACCTAGATGGCCAGGAAGTGACTTGTGCGGGAGCCTT
 tValProValAspSerThrValHisLeuAspGlyGlnGluValThrCysArgGlyAlaLe
 310 330 350
 GGCACCTCCCAGTGGCCAGCTGGACCTGCTTGGCCTGGGCTGGTAGAGCCAGGCACCCA
 uAlaLeuProSerAlaGlnLeuAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGl
 370 390 410
 GTGTGGACCTAGAATGGTGTGCCAGAGCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCT
 nCysGlyProArgMetValCysGlnSerArgArgCysArgLysAsnAlaPheGlnGluLe
 430 450 470
 TCAGCGCTGCCTGACTGCCTGCCACAGCCACGGGGTTTGCAATAGCAACCATAACTGCCA
 uGlnArgCysLeuThrAlaCysHisSerHisGlyValCysAsnSerAsnHisAsnCysHi
 490 510 530
 CTGTGCTCCAGGCTGGGCTCCACCCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGA
 sCysAlaProGlyTrpAlaProProPheCysAspLysProGlyPheGlyGlySerMetAs
 550 570 590
 CAGTGGCCCTGTGCAGGCTGAAAACCATGACACCTTCCTGTGCGCCATGCTCCTCAGCGT
 pSerGlyProValGlnAlaGluAsnHisAspThrPheLeuLeuAlaMetLeuLeuSerVa
 610 630 650
 CCTGCTGCCTCTGCTCCCAGGGGCGGCTGGCCTGGTGTGCTACCGACTCCCAGGAGC
 lLeuLeuProLeuLeuProGlyAlaGlyLeuAlaTrpCysCysTyrArgLeuProGlyAl
 670 690 710

FIG. 42

155/157

CCATCTGCAGCGATGCAGCTGGGGCTGCAGAAGGGACCCCTGCGTGCAGTGGCCCCAAAGA
aHisLeuGlnArgCysSerTrpGlyCysArgArgAspProAlaCysSerGlyProLysAs

730

750

770

TGGCCACACAGGGACCACCCCTGGGCGGCGTTCACCCCATGGAGTTGGGCCCCACAGC
pGlyProHisArgAspHisProLeuGlyGlyValHisProMetGluLeuGlyProThrAl

790

810

830

CACTGGACAGCCCTGGCCCTGGACCCTGAGAACTCTCATGAGCCCAGCAGCCACCCTGA
aThrGlyGlnProTrpProLeuAspProGluAsnSerHisGluProSerSerHisProGl

850

870

890

GAAGCCTCTGCCAGCAGTCTCGCCTGACCCCCAAGATCAAGTCCAGATGCCAAGATCCTG
uLysProLeuProAlaValSerProAspProGlnAspGlnValGlnMetProArgSerCy

910

930

950

CCTCTGGTGAGAGGTAGCTCCTAAATGAACAGATTTAAAGACAGGTGGCCACTGACAGC
sLeuTrpEnd

970

990

1010

CACTCCAGGAACCTTGAAGTGCAGGGGCAGGCCAGTGAATCACCGGACCTCCAGCACCTG

1030

1050

1070

CAGGCAGCTTGGAAGTTTCTTCCCCGAGTGGAGCTTCGACCCACCCACTCCAGGAACCCA

1090

1110

1130

GAGCCACATTAGAAGTTCCTGAGGGCTGGAGAACACTGCTGGGCACACTCTCCAGCTCAA

1150

TAAACCATCAGTCC

rt772

FIG. 42

156/157

10 30 50
CAGTGGCTACTGCTGGGATGGCGCATGTCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

70 90 110
GGGGCCTGATGGCCAGGAAGTGACTTGTCGGGGAGCCTTGGCACTCCCAGTGGCCAGCT
pGlyProAspGlyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLe

130 150 170
GGACCTGCTTGGCCTGGGCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAATGGTGTG
uAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCy

190 210 230
CCAGAGCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTG
sGlnSerArgArgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCy

250 270 290
CCACAGCCACGGGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCC
sHisSerHisGlyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaPr

310 330 350
ACCCTTCTGTGACAAGCCAGGCTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGA
oProPheCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGl

370 390 410
AAACCATGACACCTTCCTGCTGGCCATGCTCCTCAGCGTCCTGCTGCCTCTGCTCCCAGG
uAsnHisAspThrPheLeuLeuAlaMetLeuLeuSerValLeuLeuProLeuLeuProGl

430 450 470
GGCCGCGCTGGCCTGGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTG
yAlaGlyLeuAlaTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTr

490 510 530
GGGCTGCAGAAGGGACCTTCGTGCAGTGGCCCCAAAGATGGCCACACAGGGACCACCC
pGlyCysArgArgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisPr

550 570 590

FIG. 43

157/157

CCTGGGCGGCGTTACCCCATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCCCT
oLeuGlyGlyValHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLe

610

630

650

GGACCCCTGAGAACTCTCATGAGCCCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTC
uAspProGluAsnSerHisGluProSerSerHisProGluLysProLeuProAlaValSe

670

690

710

GCCTGACCCCCAAGATCAAGTCCAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGCTCC
rProAspProGlnAspGlnValGlnMetProArgSerCysLeuTrpEnd

730

750

770

TAAATGAACAGATTTAAAGACAGGTGGCCACTGACAGCCACTCCAGGAAGTTGAACTGC

790

810

830

AGGGGCGAGCCAGTGAATCACCGGACCTCCAGCACCTGCAGGCAGCTTGGAAAGTTTCTT

850

870

890

CCCCGAGTGGAGCTTCGACCCACCCACTCCAGGAACCCAGAGCCACATTAGAAGTTCCTG

910

930

950

AGGGCTGGAGAACACTGCTGGGCACACTCTCCAGCTCAATAAACCATCAGTCC

FIG. 43

1/154

SEQUENCE LISTING

<110> GENOME THERAPEUTICS CORPORATION

<120> NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES AND OBESITY

<130> 2976-4039PCT

<140>

<141>

<160> 363

<170> PatentIn Ver. 2.1

<210> 1

<211> 3626

<212> DNA

<213> Homo sapiens

<400> 1

| | | | | | | |
|-------------|------------|------------|-------------|-------------|-------------|------|
| cgggcacggg | tggccgcaa | tccagcctgg | gcggagccgg | agtgcgagc | cgctgcctag | 60 |
| aggccgagga | gctcacagct | atgggctgga | ggccccggag | agctcggggg | accccgttgc | 120 |
| tgctgctgct | actactgtg | ctgctctggc | cagtgccagg | cgccgggggtg | cttcaaggac | 180 |
| atatccctgg | gcagccagtc | accccgcact | gggtcctgga | tggacaaccc | tggcgccaccg | 240 |
| tcagcctgga | ggagccggtc | togaagccag | acatggggct | ggtggccctg | gaggctgaag | 300 |
| gccaggagct | cctgcttgag | ctggagaaga | accacaggct | gctggcccca | ggatacatag | 360 |
| aaaccoccta | cggcccagat | gggcagccag | tggtgctggc | ccccaaccac | acggtgagat | 420 |
| gcttccatgg | gctctgggat | gcaccgccag | aggatcattg | ccactaccaa | gggcgagtaa | 480 |
| ggggcttccc | cgactcctgg | gtagtccctc | gcacctgtct | tgggatgagt | ggcctgatca | 540 |
| ccctcagcag | gaatgccagc | tattatctgc | gtccctggcc | accccggggc | tccaaggact | 600 |
| tctcaaccoca | cgagatcttt | cggatggagc | agctgctcac | ctggaaagga | acctgtggcc | 660 |
| acagggatcc | tgggaacaaa | gcgggcatga | ccagccttcc | tgttggtccc | cagagcaggg | 720 |
| gcaggcgaga | agcgcgcagg | acccggaagt | acctggaaact | gtacattgtg | gcagaccaca | 780 |
| ccctgttctt | gactcggcac | cgaaacttga | accacaccaa | acagcgtctc | ctggaagtgc | 840 |
| ccaactacgt | ggaccagctt | ctcaggactc | tggacattca | ggtggcgctg | accggcctgg | 900 |
| aggtgtggac | cgagcgggac | cgcagcccg | tcacgcagga | cgccaaacgc | acgctctggg | 960 |
| ccctcctgca | gtggcgccgg | gggctgtggg | cgcagcggcc | ccacgactcc | gcgcagctgc | 1020 |
| tcacggcgcc | cgcttccag | ggcgccacag | tgggcctggc | gcccgtcgag | ggcatgtgcc | 1080 |
| gcgcggagag | ctcgggaggc | gtgagcacgg | accactcgga | gctcccccac | ggcgccgcag | 1140 |
| ccaccatggc | ccatgagatc | ggccacagcc | tcggcctcag | ccacgacccc | gacggctgct | 1200 |
| gcgtggaggg | tgcggccgag | tcggaggct | gcgtcatggc | tgcggccacc | ggcaccctgt | 1260 |
| ttccgcgcgt | gttcagcgcc | tgcagccgcc | gccagctgcg | cgcttcttct | cgcagggggg | 1320 |
| gcggcgcttg | cctctccaat | gccccggacc | ccggactccc | ggtgcgcgcg | gcgctctgcg | 1380 |
| ggaaacggctt | cgtggaagcg | ggcgaggagt | gtgactgcgg | ccctggccag | gagtgcgcgcg | 1440 |
| acctctgctg | ctttgctcac | aactgctcgc | tgcgccgggg | ggcccagtgc | gcccacgggg | 1500 |
| actgctgcgt | gcgctgctg | ctgaagccgg | ctggagcgct | gtgcgcggag | gccatgggtg | 1560 |
| actgtgacct | ccctgagttt | tgcacgggca | cctcctccca | ctgtccccc | gacgtttacc | 1620 |

2/154

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| tactggacgg | ctcacectgt | gccaggggca | gtggctactg | ctgggatggc | gcattgtccca | 1680 |
| cgctggagca | gcagtggcag | cagctctggg | ggcctggctc | ccaccagct | cccagaggct | 1740 |
| gtttccaggt | ggtgaaactc | gcggggagatg | ctcatggaaa | ctgcggccag | gacagcgagg | 1800 |
| gccacttcct | gcctctgtga | gggagggatg | ccctgtgtgg | gaagctgcag | tgccaggggtg | 1860 |
| gaaagcccag | cctgctcgca | ccgcacatgg | tgccagtggg | ctctaccgtt | cacctagatg | 1920 |
| gccagggaag | gacttgtcgg | ggagccttgg | cactccccag | tgcccagctg | gacctgcttg | 1980 |
| gcctggggcct | ggtagagcca | ggcaccagct | gtggacctag | aattggtgtg | cagagcaggg | 2040 |
| gctgcaggaa | gaatgccttc | caggagcttc | agcgctgcct | gactgctgc | cacagccaag | 2100 |
| gggtttgcaa | tagcaacct | aactgccact | gtgctccagg | ctgggctcca | cccttctgtg | 2160 |
| acaagccagg | ctttggtggc | agcatggaca | gtggccctgt | gcaggctgaa | aacctagaca | 2220 |
| cttccctgct | ggccatgctc | ctcagcgctc | tgtctgctct | gctcccagg | ggcgccctgg | 2280 |
| octggtgtg | ctaccgactc | ccaggagccc | atctgcagcg | atgcagctgg | ggctgcagaa | 2340 |
| gggaccctgc | gtgcagtggc | cccaaagatg | gcccacacag | ggaccacccc | ctggggggcg | 2400 |
| ttcaccctga | ggagttgggg | cccacagcca | ctggacagcc | ctggccctctg | gacctctgag | 2460 |
| actctcatga | gcccagcagc | cacctctgag | agcctctgcc | agcagctctg | cctgaccccc | 2520 |
| aagcagatca | agtcacagatg | ccaagatcct | gcctctggtg | agaggtagct | cctaaaatga | 2580 |
| acagatttaa | agacaggtgg | ccactgacag | ccactccagg | aacttgaaat | gcaggggcag | 2640 |
| agccagttaa | tcaccggacc | tcacgcacct | gcaggcagct | tggaagtctc | ttcccagagt | 2700 |
| ggagcttcga | cccaccact | ccaggaaacc | agagccacat | tagaagtctc | tgaggggctgg | 2760 |
| agaacactgc | tgggcacact | ctccagctca | ataaacctac | agtcccagaa | gcaaagggtca | 2820 |
| cacagccctc | gacctccctc | accagtggag | gctgggtagt | gctggccatc | ccaaaagggc | 2880 |
| tctgtcctgg | gagctctggtg | tgtctccctc | atgcaatttc | cacggaccca | gctctgtgga | 2940 |
| gggcatgact | gctggccaga | agctagtgtg | cctggggccc | tatggtctga | ctgagtcac | 3000 |
| actccctgc | agcctggctg | gcctctgcaa | acaaacataa | ttttggggag | cttccctcct | 3060 |
| gtttcttccc | acctgtctct | ctcccttagg | tggttcctga | gccccacccc | ccaatcccag | 3120 |
| tgtctacact | gaggttctgg | agctcagaat | ctgacagcct | ctccccatt | ctgtgtgtgt | 3180 |
| cggggggaca | gagggaacca | tttaagaaaa | gataccaaag | tagaagtcaa | aagaaagaca | 3240 |
| tgttggctat | aggcgtggtg | gctcatgcct | ataatcccag | cactttggga | agccggggta | 3300 |
| ggaggatcac | cagaggccag | caggtccaca | ccagcctggg | caacacagca | agacaccgca | 3360 |
| tctacagaaa | aattttaaaa | ttagctgggc | gtgggtggtg | gtacctgtag | gctagctgc | 3420 |
| tcaggaggct | gaagcaggag | gatcacttga | gcctgagttc | aacactgcag | tgagctatgg | 3480 |
| tggcaccact | gcactccagc | ctgggtgaca | gagcaagacc | ctgctctcaa | aataaatttt | 3540 |
| aaaaagacat | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | 3600 |
| aaaaaaaaaa | aaaaaaaaaa | aaaaaa | | | | 3626 |

<210> 2
 <211> 227
 <212> DNA
 <213> Homo sapiens

<400> 2
 accggggcagc ggtcgggccc aatccagcct ggggggagcc ggaagtgcga gccgctgcct 60
 agaggcccgag gagctcacag ctatgggctg gaggcccccg agagctcggg ggaccccggt 120
 gctgctgctg ctactactgc tgctgctctg gccagtgcca ggccggggg tgcttcaagg 180
 acatatccct gggcagccag tcaccccgcga ctgggtcctg gatggac 227

<210> 3
 <211> 3509
 <212> DNA
 <213> Homo sapiens

<400> 3
 cagctatggg ctggaggccc cggagagctc gggggacccc gttgctgctg ctgctactac 60
 tgctgctgct ctggccagtg ccaggcccg gggtgcttca aggacatatc cctggggcagc 120
 cagtcacccc gcactgggtc ctggatggac aacctggcg caccgtcagc ctggaggagc 180
 cggctcgcga gccagacatg gggctgggtg cctggaggc tgaaggccag gagctcctgc 240

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ttgagctgga | gaagaaccac | aggctgctgg | cccaggata | catagaaacc | cactacggcc | 300 |
| cagatgggca | gccagtggtg | ctggccccc | accacacgga | tcaattgccac | taccaagggc | 360 |
| gagtaagggg | cttccccgac | tccctgggtag | tccctctgcac | ctgctctggg | atgagtgggc | 420 |
| tgatccacct | cagcaggaat | gccagctatt | atctgcgtcc | ctggccaccc | cggggctcca | 480 |
| aggacttctc | aaccacagag | atcttctcgga | tggagcagct | gctcactctg | aaaggaacct | 540 |
| tggggccacg | ggatcctggg | aacaaagcgg | gcatgaccag | ccttctctggt | gggtccccaga | 600 |
| gcagggggcag | gcgagaagcg | cgcaggaccc | ggaagtacct | ggaactgtac | attgtggcag | 660 |
| accacacacct | gttcttgact | cggcaccgaa | acttgaacca | caccaaacag | cgtctctctg | 720 |
| aagtgcacca | ctactgtgaac | cagcttctca | ggactctgga | cttcaagggt | gcgctgaccg | 780 |
| gcctggagggt | gtggaccgag | cgggaccgca | gccgcgtcac | gcaggacgcc | aacgcccacgc | 840 |
| tctgggcctt | cctgcagctg | cgccgggggg | tgtggggcgca | gcggcccccac | gactccgcgc | 900 |
| agctgctcac | gggcccgcgc | ttccaggggc | ccacagtggt | cctggcgccc | gtcgagggca | 960 |
| tgtgcgcgcg | cgagagctcg | ggaggcgtga | gcacggacca | ctcggagctc | cccatcggcg | 1020 |
| cctgagccac | catggcccat | gagatcggcc | acagcctcgg | cctcagccac | gaccccgcag | 1080 |
| gcgcgtcggt | ggaggctcgc | gccagctcgc | gaggctgcgt | catggctgcg | gccaccgggc | 1140 |
| accgctttcc | ccgcgtgttc | agcgctcgca | gccgcgcgca | gctgcgcgc | ttcttctcgca | 1200 |
| agggggggcg | cgcttgcttc | tccaatgcgc | cggaccccg | actcccggtg | cgccggcgcc | 1260 |
| tctgcgggaa | cggcttcttg | aggagcggcg | aggagtgtga | ctgcggccct | ggccaggagt | 1320 |
| gcgcgcgaact | ctgctgcttt | gctcacaact | gctcgctgcg | cccggggggc | cagtgccccc | 1380 |
| acggggactg | ctgcgtgcgc | tgccttctga | agccgctggt | agcgtctg | cgccaggcca | 1440 |
| tgggtgactg | tgacctccct | gagttttgca | cgggcacctc | ctcccactgt | ccccagacg | 1500 |
| tttaactact | ggacgggtca | ccctgtgcca | ggggcagtg | ctactctg | gatggcgcat | 1560 |
| gtcccacgct | ggagcagcag | tgccgacagc | tctggggg | tggtctccac | gactctccc | 1620 |
| aggcctgttt | ccaggtgtgt | aactctgcgc | gagatgctca | tggaaaactgc | ggccaggaca | 1680 |
| gagaggggcca | cttctctccc | tgtgcaggga | gggatgccct | gtgtgggaag | ctgcagtgc | 1740 |
| aggtgtgaaa | gccagcctg | ctcgaccgcg | acatgggtgc | agtgactctc | accgttcaac | 1800 |
| tagatggcca | ggaagtgtact | tgtcggggag | ccttggcaact | ccccagtgc | cagctggacc | 1860 |
| tgcttggcct | gggcctggta | gaccagggca | cccagtgtgg | acctagaatg | gtgtgccaga | 1920 |
| gcaggcgctg | caggaaagat | cgcttccagg | agcttccagc | ctgctgtact | gctgtccaca | 1980 |
| gccacggggg | ttgcaatagc | aaccataact | gccactgtgc | tccaggtctg | gctccacacct | 2040 |
| tctgtgacaa | gccaggcttt | gggtgcagca | tggacagtgg | cctgtgcag | gctgaaaaac | 2100 |
| atgacacctt | cctgctggcc | atgctcctca | gcgtcctgct | gcctctgtc | ccagggggcg | 2160 |
| gcctggcctg | gtgttgtctc | cgactcccag | gagcccatct | gcagcgatgc | agctggggct | 2220 |
| gcagaaggga | cctgcctg | agtggcccca | aagatggccc | acacagggac | caccccctg | 2280 |
| gcggcgttca | ccccatggag | tggggcccca | cagccactgg | acagccctg | cccttgacc | 2340 |
| ctgagaactc | tcatgagccc | agcagccacc | ctgagaagcc | tctgccagca | gtctcgctg | 2400 |
| acccccaa | tcaagtccag | atgccaaagt | cctgcctctg | gtgagaggta | gctcctaaaa | 2460 |
| tgaacagatt | taaagacagg | tggccactga | cagccactcc | aggaacttga | actgcagggg | 2520 |
| cagagccagt | gaatcacogg | acctccagca | cctgcaggca | gcttgaagt | ttctcccc | 2580 |
| agtggagctt | cgacccacct | actccaggaa | cccagagcca | cattagaagt | tccctagggg | 2640 |
| tggagaaacac | tgtctgggac | actctccagc | tcaataaaac | atcagtccca | gaagcaagg | 2700 |
| tcacacagcc | cctgacctcc | ctcacacgtg | gaggctgggt | agtgcctggc | atcccaaaag | 2760 |
| ggctctgtcc | tgggagctct | gtgtgtctcc | tacatgcaat | ttccacggac | ccagctctgt | 2820 |
| ggagggcatc | actgctggcc | agaagctagt | ggctcctggg | ccctatgggt | cgactgagtc | 2880 |
| cacactcccc | tgcagctcgg | ctggcctctg | caaaacaaaca | taattttggg | gaccttctt | 2940 |
| cctgtttctt | cccacctgt | cttctccct | agggtggtcc | tgagccccc | cccccaatcc | 3000 |
| cagtgctaca | cctgaggttc | tggagctcag | aatctgacag | cctctccccc | attctgtgtg | 3060 |
| tgtcgggggg | acagagggaa | caatttaaga | aaagatacca | agctagaagt | caaaagaaag | 3120 |
| acatgttggc | tataggcgtg | gtggctcagc | cctataatcc | cagcactttg | ggaagccggg | 3180 |
| gtaggaggat | caccagagtc | cagcaggtcc | acaccagcct | gggcaacaca | gcaagacacc | 3240 |
| gcactctacg | aaaaatttta | aaattagctg | ggcggtgggt | tgtgtacctg | taggcctagc | 3300 |
| tgtctcaggag | gctgaagcag | gaggatcact | tgagcctgag | cccaacactg | cagtgagcta | 3360 |
| tgggtggcacc | actgcactcc | actctgggtg | acagagcaag | tctcgtctct | taaaataaat | 3420 |
| tttaaaaaa | cataaaaaaa | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | 3480 |
| aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | | | | 3509 |

4/154

<211> 826

<212> PRT

<213> Homo sapiens

<400> 4

```

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu
 1           5           10           15

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln
      20           25           30

Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly
      35           40           45

Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp
      50           55           60

Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu
      65           70           75           80

Leu Glu Lys Asn His Arg Leu Leu Ala Pro Gly Tyr Ile Glu Thr His
      85           90           95

Tyr Gly Pro Asp Gly Gln Pro Val Val Leu Ala Pro Asn His Thr Val
      100          105          110

Arg Cys Phe His Gly Leu Trp Asp Ala Pro Pro Glu Asp His Cys His
      115          120          125

Tyr Gln Gly Arg Val Arg Gly Phe Pro Asp Ser Trp Val Val Leu Cys
      130          135          140

Thr Cys Ser Gly Met Ser Gly Leu Ile Thr Leu Ser Arg Asn Ala Ser
      145          150          155          160

Tyr Tyr Leu Arg Pro Trp Pro Pro Arg Gly Ser Lys Asp Phe Ser Thr
      165          170          175

His Glu Ile Phe Arg Met Glu Gln Leu Leu Thr Trp Lys Gly Thr Cys
      180          185          190

Gly His Arg Asp Pro Gly Asn Lys Ala Gly Met Thr Ser Leu Pro Gly
      195          200          205

Gly Pro Gln Ser Arg Gly Arg Arg Glu Ala Arg Arg Thr Arg Lys Tyr
      210          215          220

Leu Glu Leu Tyr Ile Val Ala Asp His Thr Leu Phe Leu Thr Arg His
      225          230          235          240

Arg Asn Leu Asn His Thr Lys Gln Arg Leu Leu Glu Val Ala Asn Tyr
      245          250          255

Val Asp Gln Leu Leu Arg Thr Leu Asp Ile Gln Val Ala Leu Thr Gly
      260          265          270

Leu Glu Val Trp Thr Glu Arg Asp Arg Ser Arg Val Thr Gln Asp Ala
      275          280          285

```

Asn Ala Thr Leu Trp Ala Phe Leu Gln Trp Arg Arg Gly Leu Trp Ala
 290 295 300
 Gln Arg Pro His Asp Ser Ala Gln Leu Leu Thr Gly Arg Ala Phe Gln
 305 310 315 320
 Gly Ala Thr Val Gly Leu Ala Pro Val Glu Gly Met Cys Arg Ala Glu
 325 330 335
 Ser Ser Gly Gly Val Ser Thr Asp His Ser Glu Leu Pro Ile Gly Ala
 340 345 350
 Ala Ala Thr Met Ala His Glu Ile Gly His Ser Leu Gly Leu Ser His
 355 360 365
 Asp Pro Asp Gly Cys Cys Val Glu Ala Ala Ala Glu Ser Gly Gly Cys
 370 375 380
 Val Met Ala Ala Ala Thr Gly His Pro Phe Pro Arg Val Phe Ser Ala
 385 390 395 400
 Cys Ser Arg Arg Gln Leu Arg Ala Phe Phe Arg Lys Gly Gly Gly Ala
 405 410 415
 Cys Leu Ser Asn Ala Pro Asp Pro Gly Leu Pro Val Pro Pro Ala Leu
 420 425 430
 Cys Gly Asn Gly Phe Val Glu Ala Gly Glu Glu Cys Asp Cys Gly Pro
 435 440 445
 Gly Gln Glu Cys Arg Asp Leu Cys Cys Phe Ala His Asn Cys Ser Leu
 450 455 460
 Arg Pro Gly Ala Gln Cys Ala His Gly Asp Cys Cys Val Arg Cys Leu
 465 470 475 480
 Leu Lys Pro Ala Gly Ala Leu Cys Arg Gln Ala Met Gly Asp Cys Asp
 485 490 495
 Leu Pro Glu Phe Cys Thr Gly Thr Ser Ser His Cys Pro Pro Asp Val
 500 505 510
 Tyr Leu Leu Asp Gly Ser Pro Cys Ala Arg Gly Ser Gly Tyr Cys Trp
 515 520 525
 Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys Gln Gln Leu Trp Gly
 530 535 540
 Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe Gln Val Val Asn Ser
 545 550 555 560
 Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp Ser Glu Gly His Phe
 565 570 575
 Leu Pro Cys Ala Gly Arg Asp Ala Leu Cys Gly Lys Leu Gln Cys Gln
 580 585 590

6/154

Gly Gly Lys Pro Ser Leu Leu Ala Pro His Met Val Pro Val Asp Ser
 595 600 605
 Thr Val His Leu Asp Gly Gln Glu Val Thr Cys Arg Gly Ala Leu Ala
 610 615 620
 Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu Gly Leu Val Glu Pro
 625 630 635 640
 Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln Ser Arg Arg Cys Arg
 645 650 655
 Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu Thr Ala Cys His Ser
 660 665 670
 His Gly Val Cys Asn Ser Asn His Asn Cys His Cys Ala Pro Gly Trp
 675 680 685
 Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly Gly Ser Met Asp Ser
 690 695 700
 Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe Leu Leu Ala Met Leu
 705 710 715 720
 Leu Ser Val Leu Leu Pro Leu Leu Pro Gly Ala Gly Leu Ala Trp Cys
 725 730 735
 Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg Cys Ser Trp Gly Cys
 740 745 750
 Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp Gly Pro His Arg Asp
 755 760 765
 His Pro Leu Gly Gly Val His Pro Met Glu Leu Gly Pro Thr Ala Thr
 770 775 780
 Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser His Glu Pro Ser Ser
 785 790 795 800
 His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro Asp Pro Gln Ala Asp
 805 810 815
 Gln Val Gln Met Pro Arg Ser Cys Leu Trp
 820 825

<210> 5
 <211> 207433
 <212> DNA
 <213> Homo sapiens

<400> 5
 gctctaataa atttgccgcc gctaatacga ctcaactatag ggagaggatc cgcggaattc 60
 ccccatgtgc catgtccacg agatgggtga acagatgaga caacatgggt gtgcagcttc 120
 tctctttttt tttttttcag acagagtctc actctgtagc ccaggctgga gtgcagtgcc 180
 gcaatttcag ctcaactgcaa cctccgcctc ccaggctcctg attcaagcag ttctcctgcc 240
 tcagcctcct gactagctgg gattacagga acgcgccact atgccagct aatttttcta 300
 ggttttagtag agacgggggt tcaccatggt gaccaggctg gtcttgaatt cctgaccttg 360

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tgatccgcc | gcctcggcct | cccgaagtgc | tgagattaca | ggcatgagcc | accgatcccg | 420 |
| gcoatgcagc | ttctcttttc | tagagttaac | aggagatgcg | gcaggctcgt | agtagccag | 480 |
| aaattctcag | acttgcactc | ttgaatttac | acgcattgtac | aaaatgcagc | tcagccctaat | 540 |
| gcctcctgtt | cagttctgtg | gtgctgccta | cttctcgtgc | ccaaattagc | cttgccgttta | 600 |
| taactaggag | gaatgacttg | gtgtcctaca | atattttggc | tcctgggctt | ctttgggtgt | 660 |
| cattctatta | ctagtgtgatt | ttttttttct | tggaagaaa | agtatctaaa | gagccagaat | 720 |
| ttactttcat | tatatctta | agttgtttaa | ttaacctatc | cttgttttca | ctttctgtgt | 780 |
| acttctttct | ttgggggtaca | aaacagtgtt | tcttttagatt | gtctattttct | aaatattatt | 840 |
| ltaaccagaa | cataccaaat | atattttcct | gcacgtttca | ttcttcattt | tacttcttaa | 900 |
| ccattgttac | gatttttttt | tttaactcgg | tctcagattt | tgctcaattg | aagttgttcc | 960 |
| agttcatgag | attttgtttt | ctttgcagct | cttgaccatg | acagatgtga | ccagcacaca | 1020 |
| gattgtttaca | gctgcacagc | caacaccaat | gactgcccact | ggtgcaatga | ccattgtgtc | 1080 |
| cccaggaacc | acagctgctc | agaaggccag | gtcagaggct | gtttcttaaa | gatttcagaa | 1140 |
| aaatcccaat | ttgtcatagg | tttagctttt | atagtgtata | tgggtataaat | aatggcccag | 1200 |
| agttactttt | caaatggggt | tctattttga | ttttattatc | cctgaggttt | tccttttagga | 1260 |
| agagatgttc | tgtatatttc | aagggtccta | ctagccctgg | gaatttgtgt | atttgttttt | 1320 |
| agaagaaggt | aggactgttg | tcctctggga | tggctgtgac | taagcaactg | atccttgggt | 1380 |
| tgaggcatct | gtagtgcac | tcactttacc | agtaccaggt | tttgataaag | ttagggttttg | 1440 |
| agagttagtc | agggtgtacg | cggcccatat | tgaactgtgg | aaatgacaga | ctttgcaaaa | 1500 |
| tctcctgtgt | ttatatcttg | gtttggcata | atctcactgc | tcactcatata | tagatttttta | 1560 |
| aatttaagtt | ataagatgac | ccaggctttt | atagtttttg | acagctttaca | agactttttt | 1620 |
| ttttgtcagt | ctcatacagt | tcataaaata | gaaaactttc | agaacttttga | agtgcagatt | 1680 |
| tgaanaagc | caagttcaac | attcccattt | tacagatgag | caggttgagg | catgggtgtt | 1740 |
| aaaagatgg | gctgcttctc | tgccagttta | aagcatgtcc | ttgagacagaa | catccaatgt | 1800 |
| cagtgttccc | cactgcagag | ctactgcaac | ttatctatct | atctgtctgt | ctgtctgtct | 1860 |
| atctctctat | ctatctatct | atctctatct | ctatctatct | gtctatctat | ttatttgcag | 1920 |
| tttgcttctc | caatttagatt | tttgtgtcta | taatatcctg | ttggctctca | ccataagttc | 1980 |
| tctctgtgga | tacaaaaagaa | aagcttttgt | ccactgtgtt | ctagatctcc | attttttagt | 2040 |
| atgagaattg | cccagaagat | aaccccatgt | actactgtaa | acagaagacc | agctgcagga | 2100 |
| gctgtgccc | ggaccagaa | tgccagtggt | agccccgaa | tcaggagtg | attgccttgc | 2160 |
| ccggtagggc | ttgcagggtc | atcttggtgt | gtgtgggtcc | atctactcag | ctgtcttccc | 2220 |
| ccaacctgt | gcagcctaag | ttgaacctag | cagaggggaa | gagctaatc | tctccattca | 2280 |
| ccccccacac | gagtattatg | ggcttttttg | tttttaacta | aaatacagtt | cttaagttatt | 2340 |
| tgttctact | gtcctttgaa | ataaagtga | acatcctttg | ctgctctgtg | gaattgagtg | 2400 |
| acagcttggt | tctgtatcac | tgagcctgct | ccttgtctct | ttcccattca | cccttgataa | 2460 |
| cctgggacta | gaccatgatg | tctcagcagt | cagctctgtg | agcaactttat | ggagagtact | 2520 |
| tctatttaac | ctactgggatt | taattgtttg | caactgtcaa | tgggccttct | ctgagaagga | 2580 |
| gaggatagat | acttctgtca | gcagcacctt | ttagggggtga | tctccagccc | tgaanaacctc | 2640 |
| aatatatcct | gcttctgagg | ttcaggatga | tgaactcagg | gcctgagacc | agccagccat | 2700 |
| gtgatataat | tggaccagg | tgggtccagaa | aggggaacct | ccttgtctgt | gcacaacatc | 2760 |
| tgagcttgtt | cagggaagtt | ggtttggaac | aggccctttt | tgaatgtccc | ttggaagttt | 2820 |
| ttgattcagc | ttgcagatgt | ggactccttt | ctatatctca | gtgtgtctca | aatttttaag | 2880 |
| tacaggaaaa | gacctgggaa | tcttgcgaaa | ggcatattct | gtgtcaagtg | tgggggtggg | 2940 |
| cctagagtg | gcattttatc | tagtcgactg | ctacactgtg | aggagcaagt | ctttgtctgt | 3000 |
| tcttaaatga | tctctttccc | atggtacctt | tcttttatct | cagtgcactg | tactgttaat | 3060 |
| gaacattggt | gaagtctcca | aagtaacttt | gttctggtga | agttgctttg | ttcttccaat | 3120 |
| gtcttccagg | aagcattcat | agctctctgt | cagtaccttg | tgtgggttca | ggaatgacac | 3180 |
| aggtagcaga | ttacaagctt | ctcttgtatg | ctatagccat | atcacttggt | ttgtttctca | 3240 |
| agaaggacat | tctcaccttg | cctttggggt | gctttgtaca | cttgatttga | ccttccacct | 3300 |
| gatgatatga | aaacagtgca | gcttttgtag | actatagatt | tgttaattcc | ttgattcaat | 3360 |
| tcactttctg | cagttttttac | cccagccctc | caatatgcac | attcattctg | ctgtcttcca | 3420 |
| cttaggattt | tagttttcta | attgttcttc | agaaggaagt | gtaccagctc | aatattggca | 3480 |
| cccacttggt | gttttcatct | aagacatagg | ataagtgaac | tcagaatatg | cttttttagga | 3540 |
| ttccggagat | atccagctga | aacattttta | aattcttgta | ttctgcattt | gttctctaatt | 3600 |
| aatgtgtcag | aggctcccac | atcctaata | agtaacctga | atttaaatga | gaaagggcat | 3660 |
| ttcgggtatc | agtaaattga | actcataata | cagtagtttt | gtctgatttc | taaaattctt | 3720 |
| ttctttctct | tttcccttta | atgaagaaa | atatctgttg | caatggctgg | catttgggtg | 3780 |
| gaaactcatg | tttgaaaatt | actactgcca | aggagaatta | tgaacaatgct | aaattgttct | 3840 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gtaggaacca | caatgccctt | ttggcttctc | ttacaacca | gaagaaggta | gaatttgtcc | 3900 |
| ttagcagct | gcgaataatg | cagtcattcc | agagcatggt | gagttataaat | cctcaaaact | 3960 |
| taagtctctg | gtttatccacc | ttcttaccaa | gggcatgact | cgagcttgca | tgtggaaggc | 4020 |
| tgtggatagt | tgtaaactgc | ttggcaagaa | ggggagtgct | ggtgaacgca | gcctgagggg | 4080 |
| ctgtgggttt | gtgctgtcag | agtctcttcc | tctttaaatt | tttaataact | tgtatatata | 4140 |
| agatctatga | ataattatat | gggggatgaa | ttgtaacatg | tatatgtgtg | cataactctg | 4200 |
| tgacatcagt | agattatttc | atacctgttt | tacctctgga | ttctgtcagg | ggagaaagag | 4260 |
| aggctactga | taattagcta | ggttggttta | agccactga | gttccttgga | gttaagggtat | 4320 |
| tataaatagt | cataagactg | tataaattacc | actaagaagt | gtacatctca | gctggatgtg | 4380 |
| gtggctcaca | cctgtaatcc | cagcactttg | ggaggctgag | gtgggtggat | cgctgaggt | 4440 |
| caggagtcca | agaccagcct | gaccaatata | gtaaaacccc | gtctctacta | aaaaatacaaa | 4500 |
| aattagccag | gcattggtggt | gtgcgcctgt | aatcccagct | actcaggagg | ccaggcgagg | 4560 |
| agaattgctt | gaacccagga | ggtggaggtg | gcagtgcagt | gagattgctg | cgcgcactc | 4620 |
| cagctctggg | gacagagact | tactcgcgtc | caaaaacaaa | aaagaacagc | aaaaaaaaga | 4680 |
| aatgtacatc | tcttgggtgc | tttcatagcc | cagccatctc | aaaagaagag | agcaccttct | 4740 |
| tgtcaagagt | ttctaagccc | agaaaggctc | aagtctctct | cttgtccacc | cagtgctctc | 4800 |
| atggggctta | tagtcaaatat | tccatgatca | catttttgtca | ttttttagtct | ggagctcataa | 4860 |
| attgtgatcc | tacogtcagt | taagttagact | catagaaaca | agctctttca | cgagtttcag | 4920 |
| ctgtggtaca | gaaacgttta | gtggaaatgt | tcttaccaa | cagggaaggag | ttgagggcaa | 4980 |
| cacttctctg | ggtcacagcct | ccttcatgtg | aaggtatgga | aatgtggcct | gggtctcggc | 5040 |
| tgctctgggc | ctctgtgtac | cacctatagg | gccattctga | gactgttcagg | aggtgcacct | 5100 |
| tattttagttt | tctccaattta | gtcccttttt | tcagtgcac | ttagatgggg | tatggacact | 5160 |
| caaacattgg | tgacatattc | ttagtgtgtt | tacctcagge | tactgtgacc | acattttgga | 5220 |
| tttcataata | ttttgatagc | tttttcagat | ttcagaatct | ccattgggtg | ctgtctctgt | 5280 |
| tgttttctct | tccatgtcca | aatctgggtc | tcttccagca | ttccatcctt | ggctggcagc | 5340 |
| tgacctttcc | ctatgtttat | cactccctca | gaaaatggat | ggcaccacgc | tttgcttatt | 5400 |
| accctggtgt | ctctaacaat | gactctcgag | tccacggaag | ttaaaagggt | tcaaccaggt | 5460 |
| ggccacagat | atacttctgg | taccctttct | ctcttcccta | gggttttctaa | ctctaaaact | 5520 |
| ttctgggtat | tctaactctg | tggtggccag | tttatgaaac | agaaattcac | agctttaatg | 5580 |
| ataagactga | cagatgagag | acaactgaa | tgtaatgtcc | ttccacagct | atacctctag | 5640 |
| tcttagccca | gttagtagag | cccagatttt | tatggaaaaa | caagaaaggga | cacctagcct | 5700 |
| aacctcttag | acagaggctg | gctggtagaa | agctggaagg | aggtgaaccc | ttgagcttag | 5760 |
| aaggagtga | ttaggatgtc | gaagacggaa | gggctttctg | tgattttaatt | agtgccccca | 5820 |
| tctgtgagat | gtagaggagg | atgattaagg | gagtggtctc | ttgagtgcag | tcgaggtaa | 5880 |
| tttgcatgtt | tggtctgcagg | ctggatggga | ggggattttt | atagttgagc | cgcaggaagg | 5940 |
| aaccaaggcc | agaccctgcg | aggccatggc | tacaatagta | atggatttga | attgtatcgt | 6000 |
| gaaggcaaa | aaattattga | aggcccttta | aaaagtattt | ttaattgtct | ttctttcttt | 6060 |
| tggaatctgt | tatttttagg | tcttttagta | ggcatgtgta | ttttttctc | tcaaaaatgg | 6120 |
| aaataggctg | ggtgtggtgg | ctcctgcctg | taatcccagc | atttttgggag | gccagtgagg | 6180 |
| gaggattgct | tgagcccagg | agttcaagac | tagcctgggc | aacacagggg | gacctcgtct | 6240 |
| ctacaaagga | agttttttta | aagaattagc | cgggcatgat | ggtggcacat | acttgtagtt | 6300 |
| ccggctactt | gggaggctga | ggcaagaggg | ttgcttgagc | ccagggggtt | gaggctgtag | 6360 |
| tgagccatga | tcatgtccct | gcactccagc | atgggcaaca | gagcgagacc | gtgctctcaa | 6420 |
| agaaagcaaa | gggaggggaa | tacagtatat | cttttgtttt | ataactacca | aaattaggaa | 6480 |
| tacttaccat | tacttggcta | aactttatat | tttgattttt | aaaacttggt | aaaaattgca | 6540 |
| atgagaagga | aatttcagga | gagcagaaga | cagactgtcc | cagggtgtac | tgtcctatta | 6600 |
| ttccctataa | aatccagctg | caggatggat | gaatggataa | agcaaatatg | gtataagttg | 6660 |
| aatatccctt | atctaaaatg | ctttggacga | ggagtgttcc | gaattttaga | atatttgcat | 6720 |
| tatactaacc | agtttaagcat | ccctaactct | gaaatccaaa | atgctctagt | gaacattttc | 6780 |
| cttcagatcc | catcatgttg | gcactcaaaa | agtttttagat | tttggaagtgt | tttggaatttc | 6840 |
| agatttaggg | tactcagcct | gtgtttgggg | gtagccatct | cttcataatg | acatttcaga | 6900 |
| acttaaatat | tgctttgtcta | taatttctgt | gaatttttga | tatatattct | tctctgagct | 6960 |
| acatttttat | cctttataaa | atggccatat | tgaagtgatg | atctatccta | atctaccatg | 7020 |
| ctcgagtcaa | gggataaaga | ggttttctct | tgtctgtggg | gtataactaa | cttgggtggt | 7080 |
| ttatctcaga | agctttgttt | ggccaagatg | ttggttatat | tcaggccagg | catgggtggc | 7140 |
| ctgcctata | atctcaacat | cttggaggcg | caagggtggg | ggatcactct | agctcaggag | 7200 |
| tttgaaacta | gccaggggcaa | catggcaaga | ctccatctcc | aaaattttaa | ttaaaaaaag | 7260 |
| atactatctg | tattcatagt | tgtgtctctt | ttgccttag | tccaagctca | ccttaacccc | 7320 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| atgggtcggc | cttcggaaga | tcaatgtgtc | ctactggtgc | tgggaagata | tgtccccatt | 7380 |
| tacaaatagt | ttaactacagt | ggatgccgtc | tgagcccagt | gatgctggat | tctgtggaat | 7440 |
| tttatcagaa | cccagtactc | ggggactgaa | ggctgcaacc | tgcatacaacc | cactcaatgg | 7500 |
| tagtgctgtg | gaaaggcctg | gtaagttcac | aggtgaatta | ggtggtatc | agagtttatt | 7560 |
| gtgagagaaa | ccataggagg | catagttcat | tgctgagatg | tgtggaagt | tcatgaaaac | 7620 |
| agatgaagta | ttgatttcaa | gcatacgaaa | aagagtataa | cccagatttc | agaaagcagaa | 7680 |
| ggaatatctc | tgggacctcg | aatagtttta | attataagca | aaactaaaaa | taactaacac | 7740 |
| tactcgaaga | aactgatatt | ctaattaaca | atgagattga | taggtttatt | gaccagaaaa | 7800 |
| agtattgaga | attgctctga | aaagcaaat | tattggtgtt | ggcagagaaa | tgtctgtggaa | 7860 |
| gaagaaacaa | aaaaggaaaa | taaaacccaa | gaaaagatta | gtaaaagcaag | caagtgcactg | 7920 |
| caggggacagt | gttcagaaaag | gtagtgtcaa | caggggagaaa | aatgatgaga | gcagttcttgt | 7980 |
| aagcgaggagc | agaagcaacc | cagaactacg | cgagagctgc | aggagagtagc | tgagcagaca | 8040 |
| gacactcgga | gtagcttccc | agctttcagt | ctctctgggc | tacattttgg | ctaactaaag | 8100 |
| gcaggcccgag | gggtagctgc | tgcagcatga | agacaactca | aaataagcct | cctcctcctg | 8160 |
| cacaggggact | cacaagcagt | cggtggaatt | gtgggttatt | ttgggaggct | ggactccatt | 8220 |
| tgcatttgtgt | ccaattttgg | gaaagttaatt | tgtctgagga | attacagagg | tataggagag | 8280 |
| aattctatagc | attgtagatt | tctaataatt | gatttctaaa | ctctcttagg | ggagctgccc | 8340 |
| agcagccttt | cagaattctg | atgaatttgt | tgaacttaat | gaatcgcgta | gaccttcttc | 8400 |
| cctccaaaat | tgcaaacatg | agattttaca | tattcaagt | gattgtaaaa | cccctaaagt | 8460 |
| tcatctctggg | tccccaaat | ctaaggggct | tacagcccta | cttttatgga | aagattcaact | 8520 |
| gtttctctcat | ctttgtcctt | aatgctcttt | gaaaagaaaa | tttatttttt | ccacaagtgt | 8580 |
| gtaaatacac | ttgactaaac | agtacttgat | gacttttatt | gttttctatt | ttctctcatt | 8640 |
| taatttgatt | catggcacat | taatgggtca | tcaccacat | ttgaaaagtc | ttgggtgggc | 8700 |
| acagtggctc | acgcctgtaa | tccagcact | ttaggaggcc | aagacgggtg | gatcacagg | 8760 |
| tcaagatatac | aagaccatcc | tgggtcaact | agtgaaaacc | catctctact | aaaaatagca | 8820 |
| aaaattagcc | agacgtgatg | gcgcaccctt | gtagtcccag | ctacttggga | ggctgaggca | 8880 |
| ggagaatcac | ttgaaccggg | gaggcgagg | ttgcagtga | ccgaaattgt | ggcactgcac | 8940 |
| tctagcctgc | agacagagt | agactccgtc | tcaaaaaaaa | aaagaaaaag | aaagaaagaa | 9000 |
| aagtctttga | ccttagcgga | catggttggg | gctctaagt | tctctctgg | gtttcttcc | 9060 |
| cagcaaaacca | cagtgtcaag | cagtgccgga | caccatgtgc | cttgaggaca | gcagtgtggag | 9120 |
| attgcaccag | cggcagctct | gagtgcatgt | ggtgcagcaa | catgaagcag | tgtgtggaat | 9180 |
| ccaattgcta | cgtggcctcc | ttcccttttg | gccagtgat | ggaattggat | acgatagaca | 9240 |
| cctgccccc | taagtgaaaa | agggagccct | aggcacttat | gcataccctc | tgtataggca | 9300 |
| acaactcagc | tcatgggtg | tgctgtcagc | ctctgaacat | tttagaatac | agactggaca | 9360 |
| tgaacctctg | tcaaacctga | ccagagactg | ccatcgagac | cttctgcctt | attgagaacc | 9420 |
| ttcatcacaga | atcaggcaca | ttgacagtaa | ataaatgtaa | gatagatcac | agagtacaga | 9480 |
| aataaactgt | ccaacttcag | tgtcatattg | ctcaatccat | gtaatatctc | catatctgaa | 9540 |
| tttccataat | tgtaaagtaa | atgctttcca | atagataatc | tctaaggctc | cttttgctt | 9600 |
| caacatcctg | ggattgagag | aggagggaag | ggtcatctct | gttatgtatt | gggcataaata | 9660 |
| ctgggctctt | tacattcatt | atctctttta | aataatcaag | acagaataat | atttttgact | 9720 |
| caagccagtt | gaatagtctg | ttaaaaaaaa | agtaaataca | gtgaattcag | atctacctgt | 9780 |
| gatagtcaat | tgcaactttt | tttttttaat | agctgaaaat | tgttccaggct | actgtacctg | 9840 |
| tagtcatctg | ttggagcaac | caggctgtgg | ctgggtgact | gataccagca | atactggcaa | 9900 |
| agggaaatgc | atagagggtt | cctataaagg | accagtgaag | atgcctctgc | aagcccctac | 9960 |
| aggaaatttc | tctccacagc | ccctgctcaa | ttccagcatg | tgtctagagg | acagcagata | 10020 |
| caactggtct | ttcatttact | gtccaggtaa | gatgccttgc | atatccaaat | tcaagtgttt | 10080 |
| cactactgat | ttatgaagaa | taaaaccttg | aaagctacgt | tgtgtatagt | taactccctg | 10140 |
| cctcagccc | cttctcttcc | tctcaatggt | tggtaacaaga | aggaatagac | cagaagctgg | 10200 |
| tccaaggcct | gacctggacc | tgctgagagt | ggtgggtggg | tcctaagaaa | ccaattctaa | 10260 |
| gaaattggcc | tttgattcag | acttgaagtg | accactcagc | aatgtgtctg | tgggtttcta | 10320 |
| gaacagttgg | gagagcttgg | gctggtgcga | agactcctca | gagattagca | gtcaagaact | 10380 |
| tctctaagag | cctgccattg | acaacagggc | tgtttgtgag | gactttgtaa | gggaaagtcc | 10440 |
| actgtaaaca | aagctaaaa | ggcagagaca | gactggggaga | aaataacctgc | actgcactga | 10500 |
| acaactgatg | atcatccaga | atatgtaaac | tcccacctc | caggggaaaaa | ataaagaagt | 10560 |
| aaatgtgaac | ccaacagaaa | aagtgggtat | tggagataaa | gaagccattc | tcagaaaaagg | 10620 |
| aaatagaaca | ggaaaaatgt | aactaacacc | aggaaccaat | tttttgtcta | ctaaactgga | 10680 |
| tcaaatcttc | cgtgttccct | ttttccata | ctaagatatg | ggggacctgc | aattctattt | 10740 |
| ttaatctgct | aggagttaac | tttttaacga | aatattttaaa | tctctgcttt | ttcatgaata | 10800 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tcacgaatat | atctcggtaaa | atgacaaccc | agagaatggg | agaaaaatatt | tgcaaaagtat | 10860 |
| atatctaata | agaatccaat | gtccagaaca | cgtaactctt | aaaactccaac | aatagaaaga | 10920 |
| caacccaatt | agaaaatgga | taaaaggattt | aaatagacat | atgtccagag | aagaaaataca | 10980 |
| aatggccaat | aagcacatga | aaagatactc | aatatcattc | atcattacca | agaaatgcaa | 11040 |
| gtcaaaagcca | caatgagata | ccactttaca | cccactgaga | tggtctgaat | caataaaaaca | 11100 |
| ggtaataaaca | agtatggaaa | ggatatgttag | aaattggaaac | tctcatgagc | gttggcgaggt | 11160 |
| cctcaaaaag | ttaaaatatag | agtttatcata | cgggccagca | gttttactcc | taggtacata | 11220 |
| ccaagaaaa | ttgaaaacat | atgttaccca | aaaacttgta | tataaatgct | tatgcttata | 11280 |
| gcacattctt | tcatagcatc | atgctcaaa | caacattatt | cataataagc | aaaagtgaag | 11340 |
| acaagccaaa | tacctgttag | ctggcaaatg | gatgaacaaa | gtggtgtata | ttcatacagt | 11400 |
| gtaatgttat | ttggcaataa | aaaggaatga | agtactgaca | tattttctga | catgggtacc | 11460 |
| cctaataaaa | catgctatgt | aaaagaagcc | aggtgcacaa | gattttcgct | tgcatgatgc | 11520 |
| cattttatat | aaatgtccaa | agacagaag | tagatgttta | gtggtttcct | agggctgggc | 11580 |
| atggtaatga | agaataatag | gcattgaggt | ttctagagta | gctgcagcat | tattttctga | 11640 |
| catggattac | ctaaaaaaa | atgctatgtg | aaagaagcca | ggtacaaaag | attatgcatt | 11700 |
| gcattgatgcc | attttatatga | aatgtccaaa | ggcagaaagt | agatgttttag | tggtttcccta | 11760 |
| gggtctgggga | tggaacaag | gaataaatagg | catgaggttt | tctagagtag | ctgcagattt | 11820 |
| attttctgac | atgaattacc | ctaaaaaac | atgctatgta | aaagaagcca | ggtacaaaag | 11880 |
| atgatgcatt | gcattgatgcc | attttatatga | aatgtccaaa | ggcagaaagt | agatgttttag | 11940 |
| tggtttcccta | gggtctgggga | tggaacaag | gaataaatagg | catgagattt | tgtagagtag | 12000 |
| ctgcacacct | tatgctcttc | accaaccata | tatgaagggt | catattcttc | cacatccctg | 12060 |
| ccaatactgt | ttactatctt | tttaatgaaa | actgttctag | tggtatgagaa | atggcatctc | 12120 |
| actgttctgt | tcatttctat | ttctctgata | actaatgaaa | ttgaacatca | acttcttag | 12180 |
| ttagcccttt | gtatatcttc | tttgggagaa | tatttagtca | aattctttgc | ccattttttca | 12240 |
| gttatgttgt | cttttttatt | attaggttgt | aagagttctt | tatgaattct | gaagtcctct | 12300 |
| attgggtata | ttaatttgaa | atactttctc | ccattatatg | gttttctttt | cactttcttta | 12360 |
| atatgccttt | tgaaatcaaa | agttttcagt | tttgattaa | ttccatgtat | cagtgttttta | 12420 |
| ttttatccca | tggtcttttt | ggatttgtat | ctagaaaatc | tggtgccaa | taacccagga | 12480 |
| tcacatagat | ttgctctctaa | gttttctttg | aaaagtttca | agattttgtg | tgtttacatta | 12540 |
| ggctcctgat | ccattttgag | ttaatttttg | tgtatgggtg | gaggtagggg | tccaaactct | 12600 |
| ctttttgcct | gtggatgtg | tcctgcacca | tttgttgaaa | agattttttt | ttttttttta | 12660 |
| ccattgaatt | ttcatggcac | atttgtgtga | gatcagttga | ctgtgaaatc | aggaactttt | 12720 |
| ttctaagctc | tcattttttg | gcagttgacc | aatgtttctc | cttttgccaa | tattacaactg | 12780 |
| ttttgatctg | tgtggctttg | tataagtttt | aaaattggga | aggctaagtc | ctccaccttt | 12840 |
| gttcttattt | tgcaagactg | ttacagctat | tctgggtttc | ttgcatgttc | atattaattt | 12900 |
| taggatgagc | ttgccatttt | ctgcacaac | aaaaagccaa | ctgggtttgat | aagggtttgca | 12960 |
| ttgaatctac | tgaccaattt | gggtgtgatt | gcctgtttgt | ttgtttgttt | gtttgtttgt | 13020 |
| ttgtttgtg | agagagggtc | tcactctgtc | accaggtctg | gagcgcagtg | gtgctatctt | 13080 |
| ggctcaactg | agctcccgtc | ttcccaggct | caagtgtatt | tccagcctca | gcctctcgag | 13140 |
| taactggtag | tacaggcatg | agctaccaac | accagctaa | tttttatatt | tttttagaga | 13200 |
| caggggtttc | ccatgttgcc | caggctgtgt | ttgaactcct | gagctcaaa | caactctctt | 13260 |
| ggcctttgoc | tcctcagttc | tgggattaca | gggtgtgagc | accacgcctg | ggcccccttg | 13320 |
| catttttaaca | aggggttaac | ttctagcctg | tgaacctgga | atgtcacaatt | attacctctt | 13380 |
| ttactttctt | ttggcattgt | tttatagttt | tcagttgtaca | tcattgtctt | gttctttttt | 13440 |
| ttaggggaaa | ataattcagt | cttttataaa | tatgatattt | gttttagttt | ttttatacat | 13500 |
| ggcctttatt | gggttgagaa | agtttctctc | tattctagtt | tttttagggg | tttgtaatga | 13560 |
| atgagtgtca | gattttgtca | aacctttttt | ctgtgtctac | ttagatgtac | acgaagtttt | 13620 |
| gttctttatt | aatactatgt | attacattaa | ttaagaatgt | taaacagatt | ttgcatctct | 13680 |
| gggataaagc | ccatttagtc | atgatgtata | cttttttatt | atgatgtata | tttttaattct | 13740 |
| gctgttggtg | tctatttgtc | agttggggat | tttgtgtgta | cattctatgag | ggggatataa | 13800 |
| gtttgtgggt | ttcttctttt | gtgattacta | tctgggtttc | atagcagggt | agtaactgggc | 13860 |
| tcctagagta | agttggaacg | tattctctct | cctctattta | ctgggaagagt | tcgtcaagga | 13920 |
| ttagcattaa | ttcttcttta | gttgttgatt | gaattcacca | gtgaaaccaa | atgggcctgg | 13980 |
| cccttttctt | tacaggaaaa | tatttttaata | tttaattcaat | ctgtttgtta | tagactctatt | 14040 |
| cagatttttt | ttttctaggt | gagtcagttt | tggtaattctg | tttttctag | aatttgtcta | 14100 |
| ttctcagtaa | tcaaatttgt | taacatacag | ttcgtagtat | ttcttcttgt | tttttttttt | 14160 |
| ttttttttga | gacagagctc | cactctttca | cccagttctg | agtgacgtga | cgtgatctct | 14220 |
| gcttactgca | acctccgctc | ccagggttca | agtgattctc | ctgcctcagc | cccccaagta | 14280 |

| | | | | | | |
|-------------|-------------|------------|-------------|-------------|-------------|-------|
| gctgggacta | caggcgacag | ccaccactcc | tagctaattt | tttattttgt | tttgttttgc | 14340 |
| tttttttagt | agagacaggg | ttttaccatg | tagaccagga | tgggtctgat | ctcctgacct | 14400 |
| cgtgatctgc | ctgccttggc | ctcccaaagt | gctgggatta | caggcgtag | ccaccacgcc | 14460 |
| cgcccgagtc | atagatattt | cttcgaggtt | gctcttttag | aatatgtgat | ctcctgttta | 14520 |
| acctgttaac | gatttttttt | ctgtcttaat | ttggttttat | tattctttta | catttctgca | 14580 |
| gggctctaaa | tattgtgagg | gttttttttt | ttaagagctt | gtctctttat | atgtgtgatg | 14640 |
| caataattta | agcctatttg | aatgtactaa | aattcgtatt | tctctattct | tcatatttct | 14700 |
| tgcattaaac | cttccttcag | gattttgttg | tgtgcgtggt | tatcctgggt | gtattcggag | 14760 |
| attgagagtt | ttcttatact | gggttaactc | caaggttaat | ttgtattaat | aactaaatgt | 14820 |
| gtagatctgt | caatattggg | cagtggttgg | gtgtgctgag | ctgttttgaa | agtgtgggtat | 14880 |
| gatgtcctgg | aaaagctagc | aagccagcct | caggacactc | cagctctggg | ccacttcgac | 14940 |
| tgttgttgac | tgggtgtcgc | atctccctca | ttggagaag | gaaggagcaa | tggacctggg | 15000 |
| tggctgggtg | caattagcat | gttcattggg | gtggcaagtg | cctcccagcc | ttgggtggta | 15060 |
| gggccaccta | aagatagctg | gcacattgcc | ttgagttttc | tttcatttct | gcttgcttgc | 15120 |
| ttgcttttat | ttttgagatg | gagtttctac | cttgttgccc | aggetggagt | gcaatgggtg | 15180 |
| ttctctggct | cactgcacac | tctgctccc | gggttcaagc | ggttctctcg | ctcagcctc | 15240 |
| ccaagttagt | gggatatacg | gcacctgcc | ccacctctgg | ctaatttttt | gtatttttag | 15300 |
| tagagacagg | gtttcactgt | gttggcgagg | ctgatctcga | actcctgacc | tcaaggtgatc | 15360 |
| cacctacctt | ggcctcccaa | agtgtatggg | ttacaggcat | gagccaccac | accagctga | 15420 |
| gttttctttc | aaacactcaa | atactaaacg | gtgcataaac | agaacaagta | caaggctatg | 15480 |
| gaaggttgcc | aagatggtga | aaaactgcac | cacactgatc | aaccaaggca | gatctgaaga | 15540 |
| atcagatttt | gcttacagtg | caggttgtgc | actatgtaaa | tgtttccccc | atcctattcc | 15600 |
| cccttagagt | tctgcagtgc | acagcctgtg | cagctgtcct | gagcagtcct | cagaggtcca | 15660 |
| agcttcatat | acctttctag | ttgtaaagcc | ctatatccat | taggagtctt | agttaatccc | 15720 |
| cttgttgaaa | tgtttatcog | taaagatggg | tatttttcta | ataatgaaaa | atgtaaagca | 15780 |
| aactaaatat | tagaataata | gggaatgggt | aaataactgt | taagattcat | caaaactcatt | 15840 |
| ctaccatttt | aagtttttgt | tataattgat | gaaataaggt | ggtaaaaatg | ttgtgatcta | 15900 |
| agttaagaaa | gggtgaatatt | aatatgaaca | taattgtgca | aatattgttc | ttctggtata | 15960 |
| gaaaaaatcc | aaagaaaatg | atatcaaatg | tgaacagtg | tctctgagtg | gttggattgt | 16020 |
| atatgatctt | tccccatcct | tctatttttt | aagtatactg | tcatgagcat | acgttatttt | 16080 |
| ttagtgaata | acacacaaaa | atatttttta | atatcactat | ggggtttcac | cttgcgtgta | 16140 |
| tacattttca | attttctcgg | cttagctctt | ctttgatatt | ttttatttta | tttctcaaat | 16200 |
| tagtttttag | tgggcccagt | ttgtgtgatt | atcaaaagata | aattaactac | agagatagtt | 16260 |
| gcagataaat | cttattgaac | ttaactacat | gcagttcttc | catggtacag | accctctgaa | 16320 |
| actctttttc | ttgcagcttg | ccaatgcac | ggccacagta | aatgcaccaa | tcagagcatc | 16380 |
| tgtgagaagt | gtgagaacct | gaccacaggc | aagcactcgc | agacactgcat | atctggcttc | 16440 |
| tacgggtgatc | ccaccaatgg | agggaaatgt | cagcgttaagt | caaatttggt | aggtttactc | 16500 |
| atggcacaatc | gggtgtggac | acagacttct | tatttgactt | gaatatctag | gaggaaaaaa | 16560 |
| gccactttgt | tttggatacc | acatttctta | ttaaatcata | ctggtcagaa | gctcagctga | 16620 |
| gctggaagaa | agaaccaagg | tttcagtgt | gctggttaaa | gcaacaagcc | aaaaatgtta | 16680 |
| gagtcaagcc | tttgaccaca | ttctgcagtg | gtataagcaa | gggcccacaa | ccagagaaag | 16740 |
| tgccaaacttc | ccctgtttcc | ttctctccca | ttgaacaggg | cacaggtcaa | tagtcagagg | 16800 |
| gatcagctca | gacatgggag | tcttggggct | catctcactg | agggccggac | aaaaggttcc | 16860 |
| agcagtatat | ggacactaag | gttgtgggaa | gggcaagggt | taagggtcag | gactgggaaa | 16920 |
| gctctggggt | ggagtccag | ttaggaaagt | gtcttctgaa | ttccccctcc | tccccccatc | 16980 |
| aggagggcctt | gggtcctgaa | gaagacctca | gcccaaggcc | tcagctaggg | agcctgggaa | 17040 |
| taagggtctc | agggttagga | atgtgaacat | gtgaaggagc | gtgaccctgc | aggttagacct | 17100 |
| gagtccttaa | ctgcacctct | ctcagagccc | cacacttagg | ccacctctac | aggtcatgct | 17160 |
| gtgccaggta | aagcctgtgg | catctgcctg | ggtcaggccc | tgaaaaaatc | cacacgggct | 17220 |
| ttttaaccag | gagccagact | cctcaactct | tatagtgtat | tctaagggtt | taaaatgtcc | 17280 |
| cgtgttggct | gggtgttagt | gctcagcctc | gtaatccacg | acttttggga | gggccaaggc | 17340 |
| agggtgatca | cttgaggtta | aggagtttga | gaccagcctg | accaacatgg | tgaaccccca | 17400 |
| tctctactaa | aaatagaaaa | attagctggg | catggtggcg | ggggcctgta | atccaagcca | 17460 |
| ctcaggagcg | tgaggcatga | gaattgcttg | aacctgggag | ggggaggttg | cagtgagcta | 17520 |
| agatcacgcc | actgcactcc | agcctgggca | acagagttag | actcagctct | aaaaaaaaaa | 17580 |
| aaaaagaaaa | aaaaaaaaagt | cctgtgtttc | cccatttatg | tcaaatagaa | gctctcgaag | 17640 |
| aataaggacat | tttattatga | gaacaaaaat | tatgacaagt | gatatttaag | aactcgtttt | 17700 |
| tttattctag | tttccagcca | aaaaacataa | tgagtactat | agtcgaaaga | cttttcaagg | 17760 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ttctgagcag | gaaaagtaaa | caacataggg | aaatctctac | ctactcccca | gatcccaact | 17820 |
| ccaaactcct | tggcatggcc | cttcagaatc | tgatctgata | caagaccagc | caggagagag | 17880 |
| gcagcgttct | agctgtgtgt | cocttactct | coctgggggt | cagtttctct | ctaataatag | 17940 |
| ggctggttgt | aagatgaagg | cagagcagca | catgacacat | gttccacagg | atactagtca | 18000 |
| atacttagta | gatgtgggca | attcttatta | cocttctaata | gccattctcc | actactccta | 18060 |
| agtaaaagtt | cttgctcaga | ttaaaaaggta | tttggtgaaa | atgttttccc | cactctttgt | 18120 |
| gaatagaact | aatccgttag | acagttaacc | gagtaacctaa | cagagtgggt | gagtgtaaca | 18180 |
| ggcattgtgc | tgttaccac | tcaagacaca | ggtgctttta | ttatccccaa | gatcaagtaa | 18240 |
| actgccagg | gtctcacagt | gagatgtggc | caagctaaga | accacgccca | gtctgtgctg | 18300 |
| accttaatta | cttggctggt | tgctgttcca | gatgccatca | taccacatc | agcagctgct | 18360 |
| atctagaagc | atcacatttt | octctgtgag | atctacagg | cctgagta | atttgcctta | 18420 |
| tttttcagag | tcctaatacca | cggtagaaag | cgtggtgttt | atagaatact | gcagatggca | 18480 |
| ccaagtcttt | gatgttttct | tctttaaatt | gtatagtct | taggtaacga | tagtagccat | 18540 |
| gattttctga | atgctttactg | tgttttgagac | atcatatata | tatcatcaca | aacaccccta | 18600 |
| tgagataggt | actgttttta | tcccatacat | atggacaagg | aagctgaagg | tcagaaaggt | 18660 |
| taagtagcat | tcccagagtg | acagaagagc | caagtagagg | acgctgaatt | caaacccagg | 18720 |
| cagatggagc | taaggccttt | gcttttcaact | atacattaca | caacctcccc | cagttctcaac | 18780 |
| agagccaaag | gtcagacact | catgatctgt | acttcagcat | catctctgtt | aatctctgca | 18840 |
| ctattctgtt | ttatagtttt | ctttgggtga | ttccaaagga | tttattttta | aagcatacct | 18900 |
| ctctccagct | gaatgccttt | cattttattca | ccagcaaaac | agggatataa | gtgaaaaggt | 18960 |
| gttgaccaaa | aaggctttca | cttttttcaa | ctggagtata | atatttatca | cggtctgtat | 19020 |
| tacgtttggat | gataaaaagg | gagatgttga | gttggccaga | tgaggagagt | gggaagagct | 19080 |
| attttctgta | aaggcaaaag | aaaaggtgaa | gcaaggggtg | catgtcctga | gggcccttaa | 19140 |
| agcgtgacc | ttggcattgg | atcctgggtc | tgagccctgc | ggtgatcca | aatgtgacca | 19200 |
| tctgtggaga | ggcacctcag | ccaagaccag | coctccttca | gaggacagtc | acctccagag | 19260 |
| cacctcccta | ggggcagcag | ctatgtctgt | gtcccacac | ccagcactcg | gcattctaga | 19320 |
| gatgcttcac | aaatatattat | tgaatgaatc | aaagaatcat | agcagtga | aagagagtct | 19380 |
| attgaaaaga | tcaaaaatagt | cattgtcttca | gaaggcagtg | ggtaccagca | ttcactttcc | 19440 |
| tctctgtcct | gttttgggtc | ctttaatgtt | acctatacat | tattatcatt | agctgaaaaa | 19500 |
| tcatgaatct | taccatctga | atgctcgtac | tttaatctgt | ctttctact | gaactcaatt | 19560 |
| cagataaaat | tgcttgtttt | ggaaaaagtc | tccaatagtc | agaatttttt | aagccagttg | 19620 |
| tgacctctgt | tcctttttct | tctctgcagc | atgcaagtg | aatgggcacg | cgtctctgtg | 19680 |
| caacaccaac | acggggcaagt | gcttctgcac | caccaagggc | gtcaaggggg | acgagtgc | 19740 |
| gctgtgagta | ccatactccc | tggaccacca | gggaggacca | agagctgtg | cagctgcctg | 19800 |
| aacccccacc | tgagagccac | coacttccct | gtgtcttgtt | gctgtgggct | ctgagggatc | 19860 |
| cctgggttga | ttagtttgaa | attttgccca | ttctatttca | gacaggtcag | tccccaaaa | 19920 |
| gaggaggtcg | tcgagtttag | agcaatctct | taatggctct | tgaattcaca | ttttgtttta | 19980 |
| aatgatactg | acatttctctg | gggtgtccat | ttggagtagt | catttttaact | tcagcaacta | 20040 |
| cttgattttt | gtcatgtcaa | gagattatac | tctttcccaa | agagtagaga | tgggaagagt | 20100 |
| gggtgttttg | tgatggcatt | tatttggcgt | ttggttgtct | gatttggtaa | atgatcctct | 20160 |
| tacotttaac | atttcccatt | actctagcat | ttccttgtt | gaagcatgtc | aggcatcttc | 20220 |
| ctggagaggg | tttgaagtc | ccatccatga | ggatgcaggt | gcagcagtc | caggcatgtt | 20280 |
| taaggtttca | gacctggccc | tgcctccact | gagcaagggt | ctcttgaggc | aaatcactgt | 20340 |
| tccccccatc | tgtcagtagt | gggactagca | gtaattgttt | ttgggtgtct | ggttaggttga | 20400 |
| aacagggaag | atagttctct | aatcaatgta | taccgggctg | tctcactggg | aaacctcata | 20460 |
| aaatgcattg | atccgcattt | actcatctct | aagaaaaaac | ttaaaaagtg | tttagtgtca | 20520 |
| aaacactgag | ctgattaact | attgtaaaca | tcattatttc | ttaaaaatac | gtagtataa | 20580 |
| tgccacaggt | gattgagaca | tcattgtctg | tgctactcag | cagggttttg | tctgtctact | 20640 |
| ttttagttta | tatctgcata | tatgtttact | agcttcatgt | tgccccaaag | ccaatatgtg | 20700 |
| aagaaaagtc | taatccaagg | ttaaaattaa | cttttaggtc | cacattccct | taaacagaaa | 20760 |
| aatgtcatcg | agcaaaatta | atcactctta | coctcagtta | gctgatgcaa | caagaacatg | 20820 |
| agbtggcatc | aagtaagact | tccaagttcc | ttcgagatcg | cagatgtagg | cagggtctgc | 20880 |
| tcactcacct | ggctcacagg | acacagcata | octgctggg | aaaggccatt | accacttccc | 20940 |
| cgcctttccc | tcttcatgca | tagctgttgt | gtttacctaa | atgcttcatg | aagtgggagg | 21000 |
| ctgggttttg | ctgagtgtta | tagatgactct | gtatggggaa | attgttttct | gagttagaaca | 21060 |
| tatttttttt | ctacagactg | gtgaacactt | gttccagggt | ttaaaaaaac | agtgatctct | 21120 |
| ggtatttagt | cttctctcac | ttgtgtttaa | agaagagaat | tagtttatcc | agggaacata | 21180 |
| ggaaagaagt | gagaaattag | tatttgaaaa | atgttttgac | tgcactttta | gaagaatatg | 21240 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tagtccacac | aaaattgaaa | atgttttttca | ttatagtaaa | taggttttatt | acatttggat | 21300 |
| ttctctcatt | gctttggcaa | atgcttgtca | attgtctcca | tatagttggt | gtgtgaactg | 21360 |
| ctgagcagtc | agtattgaa | cgttcatgca | tgtctcttct | agactgctcg | agctgagtta | 21420 |
| tgggtgaagga | tgcaattata | atggctccag | actatctgta | ctttatataa | aaggatctgt | 21480 |
| ttggttttaa | aatgaattca | gtttctgttt | taaatagcag | tataaaatag | tcttttttcc | 21540 |
| ccccagcatg | tgaggtagaa | aattcgatacc | aaggaaaacc | tctcagagga | acatgttatt | 21600 |
| gtaagtgggt | ttgcaattct | tattctctaga | agcaaaagtag | ttcagtaaaa | cttcatgtgt | 21660 |
| taaacgggtt | tgagaatagt | aagtgtctata | agactatagc | agccaccaat | gaagtgttcc | 21720 |
| cagacttgat | atgtttacat | tctgtttaagt | ttactacata | tagggagcact | gttttaagct | 21780 |
| gttttaattg | tgtttgggtt | taacgttaat | gtgtccatag | cagatagcag | gagagagtag | 21840 |
| agaggcatgc | atctttgtct | atccacattt | atgttctctt | aaaactttac | tttatttgcc | 21900 |
| attacctagt | tggggctcat | atatttctgt | tttttaggatg | tagatcaaaa | acagaattct | 21960 |
| tacaatatgg | ttgaactttt | tgctattctg | tctggaaactg | atttgtgtta | accaccttga | 22020 |
| ggtgaaaggt | gagtcctgaca | aggtgaccat | gtttttatgg | gtaaatgtgt | tttctcttta | 22080 |
| tggggagaatc | cacatggtag | accagagtag | gggaccagaa | aaaaggagtt | aatggttatgg | 22140 |
| catatccatt | gttaattatat | acctgtctgta | ctgggttgaa | taaacctccc | tcaaaattca | 22200 |
| tgtccacttg | gaacctgtga | atatgacttc | atataatat | atataaaact | tctaaaaagaa | 22260 |
| aacaagagaa | aaacatcttg | accttgggtg | gggaaaagtt | gttagatagg | attctaaaaag | 22320 |
| cataaattcat | aaataaaaaag | tagactaatg | aaactatata | aaaattttaa | actttagtaga | 22380 |
| aaaaaacatc | caaaacataa | aaagccacag | actagagaaa | atactgcgac | atcatcatat | 22440 |
| atttgataag | gagtttatat | ccagaataca | tgaggaaactc | ttacaactga | ataagatgac | 22500 |
| aattccagtt | aaaaacaggg | aaagatttga | aaagacattt | catgaaaagaa | gatattacag | 22560 |
| tggccaataa | gcacattaaa | aacgacttaa | catcattcgt | tatttggaaat | ggagatttta | 22620 |
| aaaccacagt | cagatacacg | octaaagtga | ttagaatcca | acgctaagtgc | catgggctttt | 22680 |
| tagaagacag | tggttaattct | atgtctgtct | ctgcattcag | tctgttgcag | tacatctttt | 22740 |
| tggtttaaca | catgaaaaaa | acctggcctt | accaggcatg | acgttggaag | agggtagtat | 22800 |
| gatacccttt | taataagcctt | ttcagataat | tatggacgtt | acttgatatt | atgctgaaac | 22860 |
| tggacaactg | gtcatgttct | taaaagcggtt | tgatgggtta | accgaaataa | aatgataact | 22920 |
| cccacagact | tgattttctt | tggaggaact | gggtacagct | gagttaaagt | gcttctctcg | 22980 |
| ttacttaatt | tataagaaaa | gcagcctgta | tctctaaaga | cctgttctat | tttgtgtgtg | 23040 |
| tttagttttt | aaatatgcat | tctcttcttt | ccatagatac | tctctctatt | gactcatcagt | 23100 |
| tcacottttg | tctatccagc | gaagatgac | gctattacac | agctatcaat | tttgtggcta | 23160 |
| ctctgacgca | agtaagattt | tttaaagctc | tctatttttg | ttttgaattt | gtatggatct | 23220 |
| ttttcttgggt | cattacggat | ggacgtactg | ccttaacagt | gctctccaga | ctggagatac | 23280 |
| cgaagtatgc | tctagaggtta | taggaaagaa | atggttagact | ctacgttatc | tctttccac | 23340 |
| ataaaaggca | aaagtgatgt | taataagatt | taccaggatc | ttagacacag | actgacattg | 23400 |
| attccacgca | tacttactct | gcctgtcagc | ccatcatggc | ctcatacaga | aaggggactc | 23460 |
| caccatcaga | gggacagatag | cagagccttg | gtttactctc | tcaagagtga | ccagagcctt | 23520 |
| aaagacactg | ctgattgaaa | tgccagtgat | gcagccccaa | tcagacagca | agggaggggga | 23580 |
| cccacatcag | acagcaaggg | aggggacccc | aatcagacag | caagggaggg | gacactctgc | 23640 |
| tcttagagtg | agttcttagc | ctcatttagga | ggcaaaacag | caaaggctta | gttgggtcca | 23700 |
| ttaaagaagt | agccaggaat | aatttaaaatt | gtttaaataat | gtagtgaataa | tgtggattttt | 23760 |
| ttttcttgct | gtcatttaacg | acctgcotta | aggttatcacc | tatgtgttaga | gttgggttaga | 23820 |
| aggaaaggcca | cactgcggga | cattttaaacc | tgaaacatac | agaacacgaa | gatgcacctg | 23880 |
| tacagtttct | tcaatgaaat | ataaagtcac | gcagtagccca | cttcagttat | taaagaaatt | 23940 |
| ttggttaaaaa | taattggtaaa | ttatttagga | acttctcttg | agatttctta | ctctccatga | 24000 |
| acatcacacaa | agccatttac | cattacaaaa | ttccatttaac | aataaaatga | acaataataa | 24060 |
| catggaaaaac | aatatggcag | taagacaag | tataattgctt | tgtaaacaag | tgttatctat | 24120 |
| aaactctgct | aaattttaat | ttaaaagtga | aacaaagtct | atagtgtgggt | atgtttacta | 24180 |
| tacaagttaa | tgaanaatagg | atttgttttt | aatattctat | ttcaaaagata | aaataagaaa | 24240 |
| aaaatgtact | agaaaaagata | cattctcaaa | agtaaatgtc | tttttaagta | aaaaataata | 24300 |
| attactttta | atgaattatt | ttatggaaaa | actataggta | ttaatatata | atttgagtgg | 24360 |
| ctgtattgac | tggaaattttt | tcttcaacga | ttttagaata | agatttttaac | aaaagtacca | 24420 |
| tatatgaaat | gtattctactg | octcataagc | aagcgtttga | agtgctcaaa | atctttcagg | 24480 |
| atatacttca | tgccatttaac | gtcattaaaa | aaataaatat | agtagaatct | ttgttaacta | 24540 |
| tcttaaccag | gtgggaagca | accatgaaag | tatttggacc | tttctggatt | tccacgttta | 24600 |
| tcttcagaca | gaatacttac | tagagttatc | caaatgcata | tgttctgtcc | tctataaaag | 24660 |
| cacaagcatt | ttaagtttat | tgattctttc | tgtggaaaaga | catatagttg | acccttgagc | 24720 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| aatacaggtt | aaaactgtga | ggatccactt | acacgtagac | ttttttcaac | caagcgtgga | 24780 |
| atgaaaatc | agtatttgca | agatgtgaaa | cctgagtata | cagagggtgg | acttttcata | 24840 |
| tgcaagggtt | ctatggggcag | actgcgggact | ggagtatgtg | tggatttggg | catgctcagg | 24900 |
| gggtccctgga | accgatgacc | cacgtatacc | aaaggatgac | tgtaattggt | agttgtgtgc | 24960 |
| tgccagcaaa | tactaaatc | aaataagtaa | acacttgagc | tgtcccttca | agatgaaggt | 25020 |
| gaggtcttat | cagtaaatga | gaaggtaaat | gctttgtgag | agaaactctc | ggtttaatga | 25080 |
| tcacattttta | aaaatagctg | tttggaaatg | tttccattgt | tgtgattttg | tgcgtattaaa | 25140 |
| atgatcaaaa | caacaccctt | aaaaatctta | ttctaaccctc | tcagaatcct | ttaaaaatga | 25200 |
| ataatttcag | tacagtcgga | tgcattctga | aaagataaaa | ataaacattt | gaattgtttg | 25260 |
| caaaaaataa | tgtttgacc | cagttgaagag | atgtactagt | caaatttcag | tttgactaat | 25320 |
| tatttaattgc | ataattttacc | taacatcacc | aacagtcacac | ttcctccact | ggcttaacag | 25380 |
| attcctcagc | aatatcttta | ttagtcat | agaccaagga | tcaaaataaa | ttgaagttaga | 25440 |
| ttagccctgt | gaactgctat | atctctaa | ggttagacaag | gttttcaaaa | ctaagaagcc | 25500 |
| atactcaatc | atattttctct | tgaaaataata | tttttagtaa | gagaaaaata | tttttcaaat | 25560 |
| catgaatata | ataaaattat | tttttaaat | aagtaacatt | cagctctata | tgcctttttt | 25620 |
| aaaggctggt | tccagtttctg | gagaaagttt | acactgtata | taatctatga | gttttagaatt | 25680 |
| atatgggttt | cagttataaa | taaatataat | tttgggcact | tgactcaagt | ttttattgggt | 25740 |
| aggagtgaac | ttcagacatt | tgaaaaacagc | tgcccagaga | atccttagga | gggtgatttc | 25800 |
| ccagcacagg | tcaggagatg | caaatcactg | tgcctgtata | ggggctcttt | taaaaggcac | 25860 |
| tttctcatgc | agttaggcat | ttgataaagc | attaaattat | gtataactta | atgggaggga | 25920 |
| ataaaatttt | attggcatat | attgcttgta | ttaaggaaaa | ttgggttagaa | attttgcata | 25980 |
| ctcttctgtg | agtttctcta | aagacatcat | aaaatgtttg | tgattaaaga | cattttagagg | 26040 |
| agtaaacattt | attgctttat | tttaaaatct | agaaattggt | ttaatcaatt | ttctaataat | 26100 |
| ttgtaccgt | catcagcaaa | acagggattt | ggacatgttc | atcaatgcct | ccaagaattt | 26160 |
| caacctcaac | atcacctggg | ctgcagttt | ctcaggtaaa | gacataccta | gagaagacc | 26220 |
| tgcaaatgaa | gggtgtggtag | attaagaagt | gtaatatagg | aattgagaaa | gcgagctcag | 26280 |
| gagacagatt | ggtttgaaac | ccaccttgc | cacttactag | ctatgagacc | ttgagcaagt | 26340 |
| acttaaatcc | ctctctaacc | cttagcatta | ttttattcat | ctgtaaagtg | aggataatga | 26400 |
| tactatcctc | ttagaattgt | tgtatagatt | aaattagggt | actactacca | gagcttgcgt | 26460 |
| tggtgtcagg | ctcagttgtg | ggttactacc | ctagcccacc | accacattt | ctgttctctg | 26520 |
| tgtggccact | ggcactacca | catccgtgct | tcgggaagtga | aaaaattcaa | taaaattcaa | 26580 |
| tgattcgttt | caataaatga | aaacatttta | aataaaatga | gatttttagta | ggtacagaga | 26640 |
| aatgtaacct | gggaattaca | tcaagctcta | aaagcacagc | tcttgcctgc | tgcccttaagt | 26700 |
| tgattcactg | aagatctact | gtatagaaaa | ctaaagaaa | taaggatga | aggccaagtg | 26760 |
| ctgtggctca | tgcctttaat | ctcaacactt | tgggagactg | aggcaggagg | attgcttgag | 26820 |
| ctcaggagtt | taagaccagc | ctgggcaaca | tagtgagacc | ctgtctacaa | aaaaaaaaaa | 26880 |
| aaaaaaaaccc | agggttagtg | actcataact | gtaatgccaa | ctactcagga | ggctgagatt | 26940 |
| agaggatcac | ttgagcccag | gggttggaga | ttgcactgag | ccatgatcgc | accattgtac | 27000 |
| tccagtggtg | gcaaacagag | caagatccca | tctctttaaa | aaaaaaaaaa | aaaaaaaaaa | 27060 |
| aaaggataat | cactacttaa | cttgataact | caacaagtag | atatgggttt | gaaatttgct | 27120 |
| cattaaattt | acttgccacc | tgtctgttagg | caagttactt | aagggttctg | agccagtttc | 27180 |
| ctcctgtata | aagtaggata | gtaaaaacac | cgtcctggca | gggcctgata | gctcatgcct | 27240 |
| gtaatcccag | cactgtggga | agccaagggtg | ggaagatcac | ctgaggtcaa | gagttttgag | 27300 |
| accagctcgg | ccaacatggt | gaaacccctgt | ctctactaaa | aatacaaaaa | tcagccaggc | 27360 |
| gtggaggcac | gtgcctgtaa | tcccagctac | tcaggaggct | gagggtaggag | aatcgcttga | 27420 |
| accttgaagg | tagagggtgc | tgtgagccga | gatcacgccca | ctgcactcca | gcttgggtga | 27480 |
| cagagtgaaa | ctccatctca | aaaaaaaaaa | ataaaaaaac | accttcccaa | gtagagtgat | 27540 |
| gtgagaaata | aatgagataa | taaatgaagt | actcaatata | gtgcttgaaa | tgtggtgaaat | 27600 |
| ggtaactata | ttttatcatt | attactatta | caatactggg | tttttaaaaa | tcaaaaaacac | 27660 |
| aaagcaatga | gattgtgca | aaataagaat | attgccttgt | gcacgccact | tacgtttatc | 27720 |
| atcttaaaac | tattgtatga | atttgagaaa | agttcagaaa | ctctcaatga | ggagggaact | 27780 |
| ttaagaaaaa | gtctgaattca | tcagagttatt | tggagaaaagg | caacatctcc | aggcatgtga | 27840 |
| aagatttgca | atgagccggg | cggtgtgaat | tgcctgtaat | ctcagcattt | tgggaggtct | 27900 |
| gggcagggtg | attacctgag | gtcaggaggt | caagaccagc | ctgaccaaca | tgtgtaaatc | 27960 |
| ccgtccctac | taaaaaatca | aaactagctg | ggtgtggtgg | ggcgtgcctg | taatccagc | 28020 |
| tacttgggag | gctgagagag | gagaattgct | tgaacctggg | aggccgaggt | tgcgtgacc | 28080 |
| caagattgca | tcattgcact | ccagcctggg | caacaaaagc | gaaattccat | ctcaaaaaaa | 28140 |
| aaaaaaaaaa | aaaaaagatt | tgccatgagt | gtctcaatga | agacgtgata | atgtgggctc | 28200 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tagtcacagg | gtctaaactca | gacatggaaa | aaagtccatt | tcattaatct | ttatcggcac | 28260 |
| ttgaattcct | ggctaaggga | gaattgtgaa | cattgaagga | ctctctggga | ataggatgga | 28320 |
| gttataccag | attaggggga | cttaaatact | gtggtagctg | gtggtagaag | ggaggactga | 28380 |
| gtgacccttt | gaacccctcc | tccctgtctac | agtgggttag | gcagtgagcg | gtacatcagc | 28440 |
| attactggca | tgggagctcg | gcgcattgcc | aaggaggtgt | aaaggggaaa | tgcaaaggaa | 28500 |
| ttgaagtggt | gtggggcaag | tgaattgcag | tgcttgttaa | taggattcta | gtggatctcg | 28560 |
| tattttcatg | atcattgtgtg | tcacctgttt | gggggtgggg | caagggtgga | agggagtgtac | 28620 |
| atggattcct | ggtaaaaacca | tttttctttc | tttctttttt | tttttttgag | acggaatttc | 28680 |
| gctcttattg | cctaggctgg | agtgcaatgg | cgcaatcttg | gctcactgca | acctctgctc | 28740 |
| cccaggttca | agcgattctc | ctgcctcagc | ctgctgagta | gctgggatta | caggcatgcg | 28800 |
| ccaccacgcc | tggctaattt | tgtattttta | gtagagacgg | ggtttctcca | tgtttgtcag | 28860 |
| gtcggctctca | aactccgaac | tcagggtgatc | caccgccttc | ggctctccaa | agtgcctggga | 28920 |
| ttacagatgt | gagccacccg | accggcgacg | taaaaccatt | ttgggtaggg | gcataggctt | 28980 |
| gtatatcagc | ctgcgccagct | ttaaatctta | tctccatttc | ttgttggtcg | tgtgacttga | 29040 |
| gggaagttaac | ttaattttctg | tgaacctcaa | tttccagtc | tataaatgaa | gataataata | 29100 |
| gggcagctgt | gaggattaaa | tgagattgag | ctcttaaagg | ctattactgg | aacacaggaa | 29160 |
| atgtttgata | aatgctattgt | tcctttattat | taatgaggcc | agattctgtgt | ctccccctag | 29220 |
| ccccccaaa | aattgtctctt | ctctttcatg | tttttctttt | aacagctgga | accocagctg | 29280 |
| gagaagagat | gcctgttgtt | tcaaaaacca | acattaagga | gtacaaagat | agtttctcta | 29340 |
| atgagaagtt | tgatttttcgc | aaccacccaa | atatcacttt | ctttgtttat | gtcagtaatt | 29400 |
| tcacctggcc | catcaaaatt | caggttaagaa | gaggcttttg | gctctatacc | tgcaaagggt | 29460 |
| gtgaatcttc | tttagtaaga | ctaaatttac | taatttggag | cttgtggtaa | atgagatgtg | 29520 |
| caatgtggct | ttgcctttgt | aacgtgtatg | gcagaggagt | gctgagcaca | tgcatgctgc | 29580 |
| acagaagatt | ggagtgtggg | tggactgtat | cactcatgaa | agacatttgc | aaaagcactg | 29640 |
| ttgaaagcaa | gttggcatgt | aacagatttg | gtcattataa | acgtattcac | ttcttcagtg | 29700 |
| agcatttgcc | atgtgaaaagc | ctgtagggtc | acacaaagaa | cttcaattct | agagtaagggt | 29760 |
| ggatgtaagt | gaaacaaaacc | cacatataac | aactgaaagc | cagagtgtgg | aaagtgaat | 29820 |
| gagatgatcc | cagaatagtct | atcaaaagtt | aaaggaggac | aaatggggag | actatgttga | 29880 |
| agaaacatcc | ccctctagtcc | agacagaggt | ggattgatcc | ctggctccta | tacaaatcta | 29940 |
| acagcctttc | caagtctcca | gttctctgtg | catgtgacct | gaagggtagt | tgttaagggcg | 30000 |
| tgtgcagctg | ctgcagtgtc | ttgttagcct | gctcctttcc | tctgttccca | ggggggcagt | 30060 |
| gtactccctc | ttgtcccgaga | cccattggccc | cattttaact | tcttatactc | atgtccctcg | 30120 |
| gggcctttcc | tcaatacctt | ctgcttctta | ccttcttcat | ttaggtgtaat | gtggagggtta | 30180 |
| gggataggtg | ggctttcaag | gactggttca | cctttaacca | tggaaagcatg | gtcactggac | 30240 |
| ggaggctggt | gctgtttgcc | aatgttcaga | agcataatca | acctcagaag | caagtcaacca | 30300 |
| caaacatatg | aaaaaagttc | aacatcactg | atcatttagag | aaatgcaaat | cagaacacca | 30360 |
| gtgagatacc | atctcacacc | agtcagaatg | tctgttatta | aaaagtcaaa | aaagaacaga | 30420 |
| tgctggcaag | gctgtggaga | aacagggaacg | cttttacact | gttggtggga | gtgtaaatta | 30480 |
| tttcaaccat | tgtggaagac | agtcctggcaa | ttcctcaaaag | acctagaggc | agaaaatacca | 30540 |
| tttgaccocag | ccatcccat | actgggtgta | tatccaaaag | aatataaact | attctgtaac | 30600 |
| aaagatacat | gcacatgtat | gttcattgca | gcactattca | cagtagcaaa | gacatagaat | 30660 |
| caacctaaat | gcccatacgt | gatagactgg | ataaagaaaa | tgtgggtacat | ataccactg | 30720 |
| gaatactatg | cagccataaaa | agcgaatgag | atcatgtctt | ttgcagggaac | atggatggag | 30780 |
| ctggaggcca | ttatcctcag | caaacctaag | caggaaacgga | aaaccaggta | ccacatgttc | 30840 |
| tcactcgtaa | gtgggagccg | agaaacacatg | gacacatggt | ggggaaacac | acacactggg | 30900 |
| gcctattgga | gagtgatggg | gagggaaggag | agcatcagga | agaaatagcta | gtggatgctg | 30960 |
| ggcttaatac | actggtgatg | agatgatgtg | tgtagcaaac | cactgtggga | acggtttacc | 31020 |
| tatgtaacaa | ccgtatacat | octgcacatg | taccocctgaa | cttaaaatt | aaacaataca | 31080 |
| aaaaagcaa | gttatcactc | atcaattagg | atgccttggg | actgtgacta | aagaacagtt | 31140 |
| gatctttcca | ttctggaat | atggaggaca | aaaactatgt | tgtagttttt | ctcaccacc | 31200 |
| tctctccttt | ttctgtcttc | catggtctaa | gatttggtaa | tcctctcact | aggttctccc | 31260 |
| tttgccctcg | cataactcct | gtgtccctcc | tcagccagtc | ttgtaagcat | cagccacgct | 31320 |
| tctattcctc | gttttctctt | gtctagtcc | acatctgctc | acaatgggtc | ggcctcttcc | 31380 |
| ctcaccacat | octgaaactg | tcctgtcata | gtggggccca | gtgatggaa | ttttcagttc | 31440 |
| octttttatc | tcatttgcag | gctttgaatc | atcatcttcc | ttgtcttcc | ggatttttct | 31500 |
| tttctcggtt | ttctggccag | gaatcctctt | cctcttctct | aaacactggg | catgtgggac | 31560 |
| cttgtctagg | attcagacct | ggctcctttc | tctcttctct | ctgtttctcg | catgtgggac | 31620 |
| tttgtgtgcc | acatctattc | tggtgattct | cagtgctgtc | tcagaccgg | cactggcagt | 31680 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tgcgtgcctgg | acagccctac | ctgggtagtc | aaagctcccc | gcgttcggtg | catattccct | 31740 |
| cttctgatct | tagcccccaga | cttccatgct | tcctttctct | gtbaaacagca | ccaccatcac | 31800 |
| cctgtgttac | aagctagggc | ctgatgatca | tttgcatcc | tggggataaac | ccaacttgag | 31860 |
| catgacgtat | tatttcccta | ctttacacgt | tgttggattc | agtttgcata | tatttagggt | 31920 |
| agaattttgt | atctatttca | tgagtaagaa | tgtctgataa | ttttcccttc | ctgtccctgt | 31980 |
| tatggttttg | atctcatggt | aaactaacct | tttaataatta | taatttgagg | agttttcaact | 32040 |
| cttctcctac | ctctggagatg | taagattaga | gttaactgaa | ccctacagta | ttggtagctc | 32100 |
| tctcctgtaa | aatcatttga | tcctagtgtc | ttatttgtgg | agatactttt | tagctgctga | 32160 |
| ttcagcttct | ttaatagtta | taggatattt | tgaatttcc | ctttatttga | ttcacttttt | 32220 |
| aatatttttt | ctagggtttt | atatcatggt | tttaaatgta | tttttaattt | tactttttaa | 32280 |
| attgatacat | aattgtacgt | atttatgagg | tacatgagat | atttttgata | atgcatacaa | 32340 |
| tgtataatga | tcaaataaga | gtaatttaga | tatccatcac | ctcaaacatt | tatcatttct | 32400 |
| tttgtttgag | tacatttcac | atcttctagc | tattttgaaa | taatgaataa | gttatttgcta | 32460 |
| actatagtca | taatgacat | tgtgcccatt | aacactagaa | cttattccct | gtaactctat | 32520 |
| tttttaacct | atcaacctct | atctccccag | ccccccagca | gctggtaacc | accattctgc | 32580 |
| actctacctc | catgagatca | gttttttttag | ctcccacatc | tgagtggagaa | aacatatcta | 32640 |
| tcctttctgtg | cctggcttgt | ttcacttaac | ataatgacct | ccagttccat | tcattgtgct | 32700 |
| gcaaaaggaca | ggattccatt | cttttttgtga | ctgaataata | tttctattga | taatatatac | 32760 |
| aacattttat | ttgtccattc | atttgttgat | ggacacgtag | gttgattcca | tgtcttcaact | 32820 |
| cttgtgaata | gtgctgtgat | aaacactttg | agcatgcagt | atctctccag | tatactgatt | 32880 |
| ttctttcttt | tggatatata | cccagcagcg | ggattgctgg | atcatgtggt | ggatctatta | 32940 |
| tgagcagctc | tcatactgtc | ttccatagt | gctgtactaa | ataatttaca | ttccccaccg | 33000 |
| cagtgaacta | gtatcgtctt | ctctgtgata | ctcgccagca | tctgttattt | tgtcttttta | 33060 |
| ataatagcca | ttttaactgg | gatgagatga | tttctcatta | tgggtttgat | ttgcatttgc | 33120 |
| ctgatggtta | ctgatgttga | gatttttttc | atatgcctgt | tggccatttg | tatgtctctt | 33180 |
| ttggacaaat | ctctattcag | cattttttgc | tcaagttttt | ttcctattga | ttcttattga | 33240 |
| gttgtttgaa | ttgctgggat | attctcggtt | taaaccctct | gttggtgga | tagtttgaa | 33300 |
| atatttttct | tcattcttata | agctgtctct | tcactctctt | gtttctcttg | cttttgtagag | 33360 |
| ctttttggct | tgatataaat | ccattttctg | atttttgctt | tttgttgctg | tgtattcgac | 33420 |
| atcttacaca | aaacatcttt | gcccagacca | atgtcctgaa | ggattttccc | aatattttct | 33480 |
| tctagttagt | ttatggtttt | aggtcctata | cttaagcctt | taatccattt | aaatttgatt | 33540 |
| tttgtatgtg | gtgagagaga | ggggcctagt | gttatttttc | tgcacatgga | tatccagttt | 33600 |
| tcccagcacc | atttattgaa | gaacctgtcc | ttccccccac | tgaatatctt | tgaactccatt | 33660 |
| attgaaaacc | agttggccgt | gaatatgtgg | atttatttct | gagttcttta | ttttgttoca | 33720 |
| ctggctctgtg | tatctgtttt | tatgtcgtta | tcattgctgt | gtgggtacta | tagcttttga | 33780 |
| gcattatttg | aagtaagggt | atatgatgcc | tccagttttg | ttctttttgc | tcagaaatgc | 33840 |
| tttggttggc | tgagtacagc | agctcgtgcc | tataatccca | gcactttggg | aggccagggc | 33900 |
| tgggtgatca | cctgaggtca | ggagttcgag | accagcctgg | ccaacatggc | aaaaccctgt | 33960 |
| ctctaataaa | aatacaaaaa | ttagccatgt | gcagtggtgg | gtgcctgtga | tcccagctac | 34020 |
| ttgggaggct | ggggcaggag | aatctottga | accggggagg | aggaggctgc | agtgagccaa | 34080 |
| gattacgccca | ctacatttcta | ccctggggcaa | acaaagcgag | actctgtctc | aaaaaaaaaa | 34140 |
| agaaaaaaaa | aaaaaagaaa | tgcctttggct | atttgaggtc | ttctgtgggt | ccatacaaat | 34200 |
| tttaggtgta | ttttttctat | ttctgtgaag | aatgtcattg | gtgtagagat | tacattggat | 34260 |
| ctgtaggtag | cttttggttag | tatggttttt | ttcacaaat | taattatttc | agttccacaa | 34320 |
| tgtggtgtct | ctttcaattt | ttttgtgacc | cttcaattt | cttaccatcag | gtttttatag | 34380 |
| ttttccttgt | aaagagggct | ttcacctcct | tgggttagatt | tatttctacg | gttgtttttg | 34440 |
| gaagttttta | aaaaatctat | cgagaaatggg | attgtctttc | tgaattcttc | ttttgtcagt | 34500 |
| togttgctcc | tatatagaaa | tgattctgat | ttttgtgtgt | tgaattttga | tccctgcaact | 34560 |
| ttactgagtt | tgtcagttct | tagtttggtg | gagcttttag | ggttttctgt | atataagatt | 34620 |
| atatcaactg | caaataggga | cacttttgact | tcctcctttc | agttttggat | cccttttaatt | 34680 |
| ctttcttttg | ctagtgtgct | ctgtgcaaat | atgttgataa | cttgttgatg | ctgtcctcta | 34740 |
| atgctcgcag | cgctctgtgt | catgaaaccc | tttgttgtat | tccaatgttg | atagcattat | 34800 |
| tcacagtaat | caaaatgtgg | aaacaaccca | aatatctatc | agtggtagaa | tgggataaac | 34860 |
| aaatgtggta | tgtatgtgtg | tgtgtgtgtg | gaaattctga | tgtgtgtgtg | tgtgtgtata | 34920 |
| ataaaatgct | attcagcatt | aaaaaggaaat | tacatcgac | aacacaaagt | aaacaaagt | 34980 |
| aacctgtaaa | acatgcctaag | tgaataagc | tagttgcaag | aggacaaatc | tttgaatgatt | 35040 |
| ccacttaaca | tgaatatctc | agagtagtac | acagatttct | agagacagaa | agtggtattg | 35100 |
| aggttaccag | cagctagggg | gagcgggaaa | tagggaatta | ctgcttcatg | gttattaaag | 35160 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| agtttccctt | tggggtgatg | aaaaagcttt | ggaagtaggt | agtgatattg | gttgccaaaa | 35220 |
| attgtgaatg | taattattgc | cactgatttg | tacacttaaa | aatgggttaa | gtgattttat | 35280 |
| gttatatctt | actacaataa | aaaaagctct | taaaaatctc | agcaagatgt | ttttgttagt | 35340 |
| atagacaagc | ttatcttgaa | atttatatga | aagggcaaaa | gttctagaat | agctataaca | 35400 |
| atttagaaaa | agaagaataa | aatgggaaga | atcaacctac | ccaatgttaa | ggcttactat | 35460 |
| ttaggtacag | taatcaagac | attgtttcag | gtagaggggt | agacacacag | atcaatggga | 35520 |
| taggtgatgc | tatagggtaga | tcccagaaat | agaaccacat | aaatatgttc | aactgtattg | 35580 |
| gatttgattt | tgttttggtg | gttttttttg | gacagggtct | cactatgatg | tcagggtctg | 35640 |
| tcttgactcc | tgggctcaag | taacctcttc | acctcagcct | cttgaataac | ttgggattca | 35700 |
| gggtgcacac | ggcacgctag | ttccagctga | attttttttt | tttttttttt | ttttttgagc | 35760 |
| aggggtccac | tttgtcgccc | aggccaaagt | gtagtgggtg | agtcacagca | cactgacgac | 35820 |
| tcacacctcc | tggtcctaag | gatcctcctg | cctcagcctc | ccaagtgtgt | gggactacag | 35880 |
| gtgcatacca | ccatgcctgg | ctaattcttt | tttttttttt | tttttttttt | tttttggtag | 35940 |
| agatgagatg | tcctctgtgt | gcctagtctg | gtctcaaaact | ctcaaaactcc | tgggctcaag | 36000 |
| tgatcctccc | gccttggtct | cccaagtgc | taagattaca | gacatgagcc | gttgcgcccc | 36060 |
| gccccagctg | gattttttgac | aaaggccctg | caaaggcaat | tcaatgaaga | aaggatagcc | 36120 |
| ttttaacgaa | tagtgctaga | gcagttggag | aaccatagga | aactaaaggga | atctctaaat | 36180 |
| ctcacatctt | ataataagaa | taacgcaata | ttgatcacag | atctaaatgt | aaacataaaa | 36240 |
| actataaaac | ttttagaaaa | aaatatagga | gaaaatcttt | gaatcccaaa | acgaaacaaa | 36300 |
| agcttttttg | actcgatact | aaaaacgtga | ttcataaaag | gaaaacttga | taaatctggac | 36360 |
| cacataaaac | caaaagtgtt | tgctctgtga | gaaacctgt | tatgaagatg | aaaagggaag | 36420 |
| ctatggactg | ggagaaaaata | tttgcaaac | atatttctgt | gaaagcactt | gaatctagaa | 36480 |
| tataataact | ctcaaaactc | aacagttaac | aaacgaacaa | ttcagttaga | aaatgggcaa | 36540 |
| aagacatgaa | tagacattta | accaaaagg | atacacatat | ggcaataaat | gacataaaag | 36600 |
| atgatggcca | tcaataacca | tttagagagg | cagggtgtgt | tcacacctga | aatcccagca | 36660 |
| ctttggggagg | ccaaggcagg | cagatccacca | cttgagggtca | ggagttcaag | accagctcca | 36720 |
| ccaacatggc | aaaaccccg | ctctactaaa | aatacaaaaa | ttagccaggc | atgggtggcac | 36780 |
| atgcctatag | tcccagctac | tcaggatact | gaggcaggag | aatcgcttga | acctggggagg | 36840 |
| cagaggttac | agtgagctga | gattgtgtga | ctgcactcca | gcoctgggcaa | cggaatgaga | 36900 |
| ctgcctctca | aaaataaaaa | aaaataacca | ttagagaaat | gcaaattgaa | agcaacaatga | 36960 |
| gataaacacta | cacacttatc | aaaatgacta | aaattaaaaa | tagtaataac | accaaatgtt | 37020 |
| ggagagagtg | gggaagagat | ggatcactca | tacattgttg | gtgggaatat | aagatggtac | 37080 |
| agtcactctg | gaaaataact | tgttctctaa | aaaactaatg | ttatacttat | catataaccc | 37140 |
| agcaattgca | ttcctagcca | tttatcctag | agaaatgaaa | accagtgctt | tcacaaatct | 37200 |
| ctgtacaccc | gtgtttatag | cacttgaaag | caatatttaa | cagctttatt | tataatagcc | 37260 |
| caaaactgga | aacaagccaa | atatctctca | acaaatgaat | gcttaataaa | atgatgctac | 37320 |
| atacacacca | tggcctaata | cccagcaaca | aaaagaaatg | aactgtgtgt | atacagaatac | 37380 |
| acaacaactt | tgataaatcc | caagcaaat | ttgctgattg | aaaaagaaaa | tcttagaat | 37440 |
| ttatatactg | catgattgca | tttatataac | atttgaaaat | gaaaatttta | gatctttat | 37500 |
| acctatggac | atattgaaata | agtttcaaat | agcttcaaat | ttctgtgact | tcttaggact | 37560 |
| ctgatctcta | gacctgttt | taaaggagct | tatctgatta | ggtcaggcct | accacacagc | 37620 |
| gagtagggaa | gatacacagc | gggtcccatc | aagaaggtgg | aaatccttgg | ggcattctta | 37680 |
| gaattctccc | tgctgtaggc | tgcatattgga | atagagcttg | caggggtttc | ggtagtttct | 37740 |
| ccgatccgaa | aaaggttgag | ctctagggtc | cagtcagtat | aatttcagtt | tgccaacatc | 37800 |
| tcttgacact | attagatat | agattatata | tatttggtag | gttattttca | tgtaaacact | 37860 |
| tgattagggg | tctagtaaat | aaaattttat | tatttactca | ttttaagttg | agattgtttt | 37920 |
| tgctactgaca | tacctctcca | tattctcaat | ctaattgaac | atcgtgcttt | gttgcttata | 37980 |
| gctgacaaat | gttgtaatact | ttttagatct | ttagtgaaaa | tcttgactta | tgctgctgac | 38040 |
| acttttccca | gtctgtgctc | ataagcaggc | ttttgtgctg | gagatggcca | tacttgttgc | 38100 |
| caatgcccag | tttttttctt | cccagttccc | tattgtctca | ccctctgaaa | gtcactgaag | 38160 |
| tggttttaga | gtctcttact | tagaggggag | tcttcccacc | accttactct | aaagaagtaat | 38220 |
| aattttaagt | tttcataact | agggagatgt | gattttcttc | ctgggtttct | gggtgctcct | 38280 |
| atcttttctg | gtacactaac | ttaaattaa | atccacttat | gaatcgtgca | ttcaggagca | 38340 |
| cttgtgaaaa | taaacgatgg | catgcccacc | atttagtttg | ccaagcttct | cttgaaaaag | 38400 |
| caggtaagtt | ctatttgaaa | cgtgtctttt | tgctgtttgg | tacaaatgac | tgcaactatc | 38460 |
| actgtgacta | ggcccggtgc | accgtgtgtt | ctcagtcctc | gcattgagagc | cttggttgaca | 38520 |
| ctgggtctct | ctgctttggg | gaccgtatct | ggctgtgcgc | tatgtggagg | gctgagctcg | 38580 |
| tgctcttctg | ggtaaatata | tcattgttct | aacaagggct | gcattgttct | ttattgtctg | 38640 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tttttcccc | gagacttagc | ttttctaggt | gtgtagcatg | gctcttaaat | acacacattt | 38700 |
| agcaactgtg | caaaattcagt | aaaaatggag | agacttattt | aaggaagaca | gtagggaagt | 38760 |
| gtctaaatag | tgccagtgtag | gccatttcagt | taaatttcag | tatggctttc | gatgcaggca | 38820 |
| aagccggttt | gccttttgatt | ctcatttttt | aaattagatt | tataaaattg | agattctcct | 38880 |
| tttttaaatg | agatttttaaa | aatttgagatt | ctcctttttt | aaatgagatt | ttaaaaaattg | 38940 |
| agattctcct | tttttaaatg | agatttttaaa | aattgagatt | ctcctttttt | aaatgagatt | 39000 |
| ttaaaaattg | agattctcct | tttttaaatg | agatttttaaa | aattgagatt | ctcatttttt | 39060 |
| taaaagcaag | taactgagga | aatctaattc | tctacttgga | aatttgtgta | agcattgggg | 39120 |
| ctttaaaaat | ggtaatacta | aactcaagtt | ccttcactct | gaaccaatcc | caggattcct | 39180 |
| ttccataatt | aggaacagat | caaaaacctat | aactgctgag | ggaatttgaa | actgactata | 39240 |
| aaaaatcttg | tagatgcagc | taaaaaacaa | acaaaaacca | cagactgaaa | agtctctgtt | 39300 |
| ctgggggggt | ggcgggtgtg | ttgggggcgg | ttgtgcctgt | tttttgattg | gttggtcctg | 39360 |
| tggttgagca | ccctccctc | tgccctgct | ttctgaaagc | tatcctagtt | gtctcccca | 39420 |
| gtaagccggg | gctcagccgg | cttcagctat | gtccatagtt | tgggggcggt | ggcggtaggt | 39480 |
| attcaagata | gaagatggtt | ggacagaaag | tgtgtgaggt | gggaaggttg | agtggtaggg | 39540 |
| ctctggctcg | ttcccgactc | tgcaacagag | gttggtactg | taactgtggag | ctaatataga | 39600 |
| agtcctcggt | ctgggtgttt | catctactgt | tttatgtaa | agggcagttg | tcacaagtaa | 39660 |
| aggaacacaca | gaatttggtat | tgcagaactgg | gcattcattt | agattttaga | gttctgttta | 39720 |
| tggtgtctaa | caggatcgct | atactgagca | agggatgtgt | gaaacttttg | ttcacttgga | 39780 |
| tgtgtctttt | ttcttataaac | agcttttattg | agacatcatt | tacataccat | acagttcgct | 39840 |
| catttaaaagt | acacaattca | gtggttttca | gtgtattcac | agaagcggtg | agccaccctg | 39900 |
| cactcacagcc | aattttagaa | tattttcacc | acttccagaa | gaagccocta | gcagtcagtc | 39960 |
| ctcattcccc | cctcttccgg | caacccccag | tctgtatttc | ctgtctctgt | ggatttgctt | 40020 |
| attctggaca | ttttatataa | attgagtcc | atgatatgct | ctcacctagc | atactgtttt | 40080 |
| caaggtttcac | ccgtgtttag | tatgtttcag | tacttattcc | ttgtcatagc | tgaataatct | 40140 |
| tccttatatg | tatataccac | gttttgttta | tccatttttc | tatcgatgga | tgcctgggtg | 40200 |
| gttttctact | tttggctatt | atgataagtgt | tcagatatac | ttttaatggt | tatatattat | 40260 |
| gtataataac | agtgcatatc | cgtaactcct | ggtgttaatt | gtctgtatct | taagagtttt | 40320 |
| aagaattata | tggtttgtct | ctaacttagga | ggccgagcgg | gcccagcgcg | gcagatccg | 40380 |
| tttagctcag | gagtttgaga | ccaggctggg | caacatagta | aaacctgtc | tctactaaag | 40440 |
| tacaaaaaat | tagctggggc | tggtgaagtg | ttcctgtagt | cccgactact | tcgaagctgt | 40500 |
| aggcacaaga | atcacctgaa | cccgaggagc | agatgttgca | gcgagccaa | atgacaccac | 40560 |
| tgtactgcag | cctgagcgac | agagcaagac | tctgtctccc | aaaaaaaaaa | aaaatttttc | 40620 |
| tttatgatga | aaaccacaag | gctgtacaag | aaccttaaat | atagtttgaa | acaaagaaag | 40680 |
| acacatgctt | tttaagtacag | tttaattctt | ttcttttctc | ttatttcatt | tttttttttt | 40740 |
| ttggagacag | agtcctcact | tgctgcccag | gctggagtac | agtggtgcaa | tcttggtctg | 40800 |
| tgcaacatc | tgctctctgg | gttcaagtga | ttctcgtgcc | tcagccttgc | gagtagctgg | 40860 |
| gactcacagga | acacttcacc | acgcccagct | aattttttgt | atttttagtag | agacaggggt | 40920 |
| ttaccatggt | gccccatca | ttctcagcaa | actatcgcaa | ggacaaaaaa | ctgaacaccg | 40980 |
| cttgtttcoa | ttcataggtg | caaaactgaac | aatgaggacg | catggacaca | ggaaggggaa | 41040 |
| catcacacac | cagggcctgt | tgtggggttg | ggggaaaggg | aagggatagc | attaggggat | 41100 |
| atacctaagt | ttaaatgacg | agtttaattgg | tgccagcacac | caacatggga | tatgtataca | 41160 |
| tatgtaaaca | acctgactgt | tggtcacatg | taccctaata | cttaaaagtat | aattaaaaaa | 41220 |
| aaaaaaaaaa | aaatgtgtta | ttcctttgtc | tcgaaagtcc | cctgcttttc | agcttttctg | 41280 |
| ttggttgtct | cttgtcagct | ctttaatgct | attgttccct | tggttttcat | ctggactcct | 41340 |
| tttccctggg | caatctcatc | cacaccatct | gggaacattg | cctttggctg | gcacagttaaa | 41400 |
| tctttgtctc | tagcagactt | cctctcctgt | atctccagag | ccaggaaacc | actagacttc | 41460 |
| cttaaaattc | attcagcaag | cagactttct | ggcccagagc | agcagctgtg | agtggaacgc | 41520 |
| aggcacctcg | ttagaagctt | gggtgttgag | gtgtgcccac | gcattccttg | agtgccatgag | 41580 |
| aggggagggg | agtgccctgt | gtcacctgag | tccaagttaa | gagagatcac | accagagctg | 41640 |
| attcctaagg | agcatggggg | cggggggaag | gggtgttcgt | ggtggggaaa | ggaatgacc | 41700 |
| agccctttca | ggaactacca | ggagtttagc | agagctgggt | acagtgcctat | gagcaagaag | 41760 |
| ctggagagct | ggggaggggc | attatgaaga | gactggagat | tgccaggtgt | gaaacttgagt | 41820 |
| cctgagtgta | ctgggaattg | ttgagatgtt | ttaggcagaa | aagtcaaccga | cgactcagtg | 41880 |
| ttagaaagac | cccactggca | gcagttggga | gtgtggattg | gaaggagcca | ggcctaagtc | 41940 |
| ctggggagac | ttaggaggtc | caggagttcca | gggagttccag | gtggaattct | tgtgtagaaa | 42000 |
| cccagggagag | gggaccattct | gatgacacac | atcccatagc | aactctcttc | ctctctctgc | 42060 |
| acctccacag | cagagtgtct | gtttccacta | tacgaagatt | tcttaattgg | aaggatcgca | 42120 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tttttatccc | tcagtgctcc | cagtgctcatg | tgctaatacag | tggagacac | tccttaaatg | 42180 |
| tatgccagat | gagtatatag | ttgttagtat | aagggtctta | cggtaggtat | cagaactttt | 42240 |
| tcttttttat | aactttttata | actttttataa | aggagccgac | acotttttctt | gatattgtca | 42300 |
| gaacaattgc | tgaccagtg | tatgggctgg | ctgaggttgg | caaccaagaa | tctgttatct | 42360 |
| cagctctacac | agcgtggagt | tagaagattg | gattgttgtt | gggttagaatt | ttgtccagag | 42420 |
| agttagaaga | tagaagaaaa | gggggggctg | aggacataaa | catccaaatt | ttacaaagaa | 42480 |
| gaaagttcag | tgctcctgtc | tatatctgaaa | tagttaagag | ggaaaatgatt | gggtctgggg | 42540 |
| aatgcttcaa | aataatatga | tcatagggaa | agtaagtgt | gtgtggatga | aacaagggtg | 42600 |
| gccttgagct | gataatttac | taaagcagga | tgatgagcgt | gtgagggtca | ttatagtgct | 42660 |
| ctttctactt | tttaaaaaagc | ttaagatttt | ccaaaatgaa | aaagaagctg | agcatggtag | 42720 |
| ctcagctctg | taatcccagc | actttgagag | cctgagacgg | aaggatcact | tgaagccagg | 42780 |
| agttogagac | cagcctgggc | aatatagtg | gaaccccatc | cctaactaaa | aagaaaaaaa | 42840 |
| attacacaat | gaaaagaata | atacaggtgg | tagaagaaga | aataaaaggt | tgagtatatg | 42900 |
| atttggcctc | ctacagacag | ctacaggagc | atgcctctgc | tgggcccggc | tgtagcatag | 42960 |
| tgcaacctgc | agctgaagtt | ctgttcgcct | ttacaggaag | tacacagcat | tgtagcgggg | 43020 |
| ttttaccagt | taacagtgga | cagctctctag | aacactcagg | ggctcagggg | aaaactccct | 43080 |
| ctccctcaaa | ccagccccc | tgctcttggtg | agaatggaaa | acggagcttc | ccagcaagcc | 43140 |
| tctcgctctc | accocagcgc | acaggtgcac | cagggtctga | gggtctctcc | ctgtggggcc | 43200 |
| tgggaagccg | cccttcagtg | ggctatggct | gcattgtctt | ctgatcttgt | agctgtgtgt | 43260 |
| gggttaacagg | cacatcgctc | ccccatatcc | cagagagtca | tgtccacttc | ctatctgtgg | 43320 |
| tggtactctt | ccatcactgt | cacctcagg | ataatggcac | agggaactaac | agaggcacta | 43380 |
| agagcagtg | tataactctg | agggaggggc | ggggctgatg | acccaaatca | ctctctgttt | 43440 |
| cagtaagagg | ttgggtaaag | tgatgcctaa | aatggataga | tcaggaatga | acaataaaaa | 43500 |
| cattaatgat | aaaaaggcaa | ctgcagaga | aactgaaaag | tcaacctagt | caaggtgggg | 43560 |
| gtccctggag | agaaagcctc | agaagagcag | aggaagagca | gggaacattg | cttttctatg | 43620 |
| ctagcctttc | agtgccctttt | gatatttctg | ctaagttatg | agttctctaa | tgatttctaa | 43680 |
| attgttaatg | agtggtgatag | aataggaatt | ctggagaatg | gtcaaaagga | attgcttagga | 43740 |
| gaaattatgg | gaaagaaaagc | acagagcagt | attgaggtgg | gaggttgagg | gcagacagat | 43800 |
| gggtgagtgt | cggttgcatg | gcaggaatga | gtggctgaat | ggcatctaac | ggccatgggg | 43860 |
| ctctgaccag | agcttgggta | tacagtttct | gtgccagcac | aaaccagcat | taggtgcgtg | 43920 |
| gtgtgtggag | ccgggtgatg | gtcatatgga | tctggtccag | agctgtggtc | ttgacataga | 43980 |
| gttccatgta | tgaacttttg | tggtgtgggt | ggttttgttt | cgttttctaa | gaaaaaaaag | 44040 |
| agagtgtgtt | tatgtcacag | ccaaactcaa | ggaggtgaca | gcacaagacc | caggttagaa | 44100 |
| tgacttagc | cagggctcac | ggaattctcc | atcctttgca | aagccaaagaa | gagtagaggc | 44160 |
| ctcaaggaga | atagggagc | acatgggctc | cttctgggct | gagctctcac | ccttggtggg | 44220 |
| gcgcctggct | caggggggct | gttggggagc | tgccggagatt | ctgccacgta | catcacatcc | 44280 |
| ttgcgttttc | tcggagcctg | ccttgaagcg | ccttgaagcg | gaaagacttc | acctcacagg | 44340 |
| gtggccagct | gtcctttctc | ttcctctccc | tgctctgtgt | gctcatgata | atggatagat | 44400 |
| agacatctcc | atggttagaaa | taatgcagcc | ctttctgcag | tccatgggtc | cacagcatat | 44460 |
| ggtcagtaga | cttagatgag | tgctacataa | agtgtaatta | gttctgcacc | acgggtgtct | 44520 |
| accttctgaa | aaggaaactc | tggtattgac | ttagaaatgc | cagttttatc | cttccgacca | 44580 |
| gaacgcgaag | cttgaatgag | gccagaggcc | aatgaaaggt | cttttctctc | ctcttttcaa | 44640 |
| atctgctgta | ttaaacagct | gtctttctca | gtcattcagt | attcattcat | tcattccatt | 44700 |
| attcaagttt | ttgagtacct | tctttgtgga | tggcacatag | tagatgcaaa | aatacaatca | 44760 |
| tttttaaaat | ggaagcatt | tgttatgtgg | tagtggaata | taagaaatgc | catgcttcca | 44820 |
| tatttttatg | gtaactctgg | actcaggttc | taccaccat | acttaattag | aaattcattc | 44880 |
| tcagcctctg | ttttggcaga | gaccattctc | cttccctcag | gaacatacag | aaacctcagg | 44940 |
| gctgaaactga | gtcagaggt | tggggaaggt | cccctgggtc | gatcttcagt | gcctgactgc | 45000 |
| tgcccccgat | tcctctggcc | ctcactcttg | agggcagcat | cctcactctt | gtgtggcaca | 45060 |
| ggcttggaca | tagaatgaca | cttaagcacc | agagcctccc | tagcttgcca | caaggacacc | 45120 |
| ctacattccc | aggcaacttc | agcaatgcaaa | acctcccctc | catctccctc | cttccctgcg | 45180 |
| ctccaggttt | gggcccctgc | tctccttttg | tttcaaagca | ctgtcctctc | ttcctctacc | 45240 |
| cttatccttc | cccatctccc | tcccactccc | caaaagagcc | aaactaatta | tggggattta | 45300 |
| accttccaat | taaaagcgag | gatttttttt | tttcaatgca | gttctcgaaa | tgcataagaa | 45360 |
| attgtccagg | gaaaggagca | caggagaggt | tttgttttgt | tttgttttgt | tttgttttgt | 45420 |
| tttgttttgt | tttgttttga | gacagttctg | ctctgtccac | caggctggag | tgagtggtgt | 45480 |
| cgatcttggc | ttaccgcaac | ctctccacct | cctgggtttg | agcaattctt | gtgcccacgc | 45540 |
| ccccaagta | gctgggatca | caggtgcaca | ccaccacgcc | agctaatttt | tgtattttta | 45600 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gtagaaatgg | ggttttgcca | tgttgccag | gctggttctg | acctcctggc | ctcaagtgat | 45660 |
| ccaccacact | cagcctccca | aagtactggg | attaggcgtg | agccactgca | tccagcctga | 45720 |
| tgtttaataa | attctctta | tcagagtttc | ctgaagtgtg | gcctatggag | actagacatca | 45780 |
| gaatcoactgg | aggagcttaa | agatatagag | acagggatag | ggcctgagat | tctgactttg | 45840 |
| ttgttaagcta | tccagagttt | gagagccatg | gcctagttct | tggaatacaa | aagtgtatct | 45900 |
| gctcttgta | gggaggtagg | aagttaacct | ttcattcatt | cttctacaac | agtaatttaa | 45960 |
| agcttaggaa | ttgttatctt | catagacaga | agaatgcagg | aatctctctc | tcaggatgtc | 46020 |
| tgctcttaca | cagccatggt | atatggtaga | ccaagtttgt | ccaacccatg | gcccgcgggc | 46080 |
| tacatgcagc | ccaggacggc | tttgatgca | gccaaacaca | attcataaac | attcttaaaa | 46140 |
| catgattttt | ttattttatt | tttttttaag | ctcatcagca | ctatcgtag | tgtagtggt | 46200 |
| agtgtatttt | attgtgtggc | aaagacaatt | cttccagtg | ggcctaggga | agccaaaaga | 46260 |
| ttggacaccc | ctatatagac | cataattatc | acacgcctg | agaggggaagc | agatcaatgg | 46320 |
| tgagggagcg | tagaggagag | cctgccagag | tttccagtg | ggccggggat | gcttaaaaac | 46380 |
| ggagggcagt | atgtgctgag | ccagaaaggga | cagggaggac | ccccatgcac | acagggcaggg | 46440 |
| aggctctctg | gtgacatctg | cogtcaggta | cagtagtaca | gtagtagagt | gagcgggaagc | 46500 |
| acttttggct | gaagtgcac | attagtgtat | gaggaagtgc | aggctccgtaa | tatcttacc | 46560 |
| caaacctg | ggctcactga | gcttccgagt | tcagagtttt | tcgaattga | tgaacacatc | 46620 |
| cagaaggccc | attggggctg | tacaccttag | acataattgat | gaaagtggag | ttgggaattgg | 46680 |
| gtgtgtgga | tcgtaggacg | tgaattctag | gtttcttgga | gtgtggagta | tggtgtctac | 46740 |
| tgtttgga | cagagaagaa | tgacattatg | agttttattt | ataaaataag | aatttctct | 46800 |
| ttgtattaaa | atagattttt | atataggaaa | ttttcattac | cttcagtatc | agaaaggccc | 46860 |
| ccaaatctat | ttcacogttt | ccataataaa | ttgtgaaaga | aagatttgat | tgggatatgt | 46920 |
| tctttaaaaa | tgaaacccaa | aagtaacttta | ataaacttta | attttagagt | taacaacatt | 46980 |
| tatagttaca | atcatgaagc | aggatgtgtg | ctgcttcatg | agccaagtgc | ataaaagaaa | 47040 |
| aaaaaaatcag | atcttactga | gttctgtgag | acagatcatt | atttttctta | ttaacatatt | 47100 |
| catttattca | gcaaatgttt | attaagtgcc | aattaatgta | tcagtggtct | ctttaggcac | 47160 |
| tggaagata | tggtcagtg | gataagggga | agtctctgtc | ttcaatcctc | acagatctta | 47220 |
| ttttccataa | gtaagtgtac | aaatacgggc | tatgaattaa | atagagcaag | ataaagagat | 47280 |
| gggggtgtgg | cttcagttgt | gcacaaatc | cccagaaatc | tgaatgtctc | atcacggagt | 47340 |
| ccttcacttg | agactctcac | gcaggatgta | aaaagcaaca | gccaggaaac | tccttagttt | 47400 |
| tcccgaagtc | ctcccttaat | ctcgtttaga | gaattatgaa | acaacttcta | tggttctatc | 47460 |
| ctttatccct | tcttcagttg | ataactgaag | ggctaaatct | gccaaagtct | cctttgcca | 47520 |
| tggttttctg | cgagcagttc | tccaacagcg | cttccactaa | aatcttgtct | tttaacaagt | 47580 |
| ggatttttaca | gttggttggtg | attgtatgac | agtcctgtgag | tagggacagc | ccaatgtatg | 47640 |
| tttgttaatc | aagtttatgc | cttttacatt | gagtaatttt | taattacctc | gtgtaatttt | 47700 |
| taacagcact | ctccttcac | ttcagaagtg | ttcagtttga | acaataagtt | tgtcatcttg | 47760 |
| tcagtcaaa | gttgaccac | aaagactgac | cgcagtgggc | gggggtattca | gcaggggttag | 47820 |
| atgtttctaa | tgactcagtg | ctttcataaa | cattttcagc | ttatgacott | tgccctctca | 47880 |
| cacaatcttg | taactatgga | gacttagatg | ataaagattc | aactaaagag | atccotttaa | 47940 |
| agtcttgggt | accagccata | ccaaaagctg | ttttatatga | ggttctttat | gattataaatt | 48000 |
| tggaacatgaa | catggagtg | tcctgcttca | gcaagaaatc | aggctaataa | ccattagcat | 48060 |
| ccagltgtgg | taaaaggcgg | ggctgcctg | aggagtgttg | tgtggcacag | ggcgcatgaa | 48120 |
| caccagataa | gtcctcagac | agtgccctta | gcattccagt | gtttttggca | tctagcaggg | 48180 |
| agtaaaagta | tctcatagat | tcttttaagt | taaacaccga | ggacacactt | aacgttttaa | 48240 |
| aagattgaaa | tccagacagg | gctgggggtg | actctaact | gtttccttct | ctctctttat | 48300 |
| gtgttataac | ttggaattct | cattccocaa | aatgaccttg | tttatataat | tctccagttc | 48360 |
| tatctaaaaa | tttttatgtga | ggaccatggc | tcatattccg | cagctttcct | tctctcaagg | 48420 |
| agctaggact | gggttatgtg | ataaagtga | gaaaagaaaa | gaaagtgggt | ttctgtttat | 48480 |
| tgatagaaat | tgatgttaact | ttgtatttct | gattataaat | ggtagccaga | gggaggacac | 48540 |
| aaaaaacttca | aattccaagc | cccataaat | attatggcct | gaatagtgtt | ggaatttatt | 48600 |
| tatttatgga | atgtcaaaaa | gaaaatagg | ccatcaaaaa | agcaaacagt | ggaagtctat | 48660 |
| tttctggccc | ctgtgagctc | ggctctaaaa | acagtgcaag | gggttggggt | tgtgtgggac | 48720 |
| taggaagaca | gatgtctctc | atttaccctc | ctgtgtttgt | catgagaatt | ttctgtactg | 48780 |
| ggattgtcac | cttggtcagt | aattgtttcg | tttatagtag | cattgtctct | ccaagagagg | 48840 |
| ctgagttccc | taaaaaaaag | caggcagcgc | accacaacat | ctgggtagta | gacagtatat | 48900 |
| tatcgggtgat | agtaaggaca | caatctctgt | atttaataca | tttgagggta | tgtcattttt | 48960 |
| cgtgggctaa | tatccctgaa | atcatgatgc | atcttatggt | cagtgactta | aaaagcat | 49020 |
| tgtcatagtt | aaattgacag | tgtttttctc | aaattatata | cagaataatg | cctttttcag | 49080 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ccaatgttat | tttagagttg | atgaaatatg | tgtcattttt | atatctcttg | tatgaagact | 49140 |
| attaactgta | taataaaatt | attttaagt | tgtatgacgt | gtgtcagatg | atgttcagtg | 49200 |
| tagcaaaacg | cgaggcacaa | ttttttgtta | aaaggacaat | atcttgcagt | cagtacacac | 49260 |
| tgagtacata | aatgttgacc | aaatgaagag | cacagaatgg | ggctgccac | ctggaggcag | 49320 |
| agccatgaat | tggtagtgag | gcaaaacagg | agagctttgt | gtgtgggttg | aggcgcaggg | 49380 |
| agggagcgcc | aggcttcagt | gaggcatcgc | agggtctcca | gggtcaaaac | gcagaaacct | 49440 |
| cagaactggg | gtgaagctta | gaaggggctg | caggaggctg | tgtcaaaagg | tactgtgatt | 49500 |
| gctatgggac | tggcaactaa | acaagtaaac | taaaggggtg | aggtaacttt | ccttaagtga | 49560 |
| aaaaagtttc | cttaactctt | ttaaagtgt | agctctgttt | ttagcttaac | tcttgaagt | 49620 |
| agaagcgtga | atccgtgggt | gtttcatgac | ctgcataatc | acactacctg | ttttctgtc | 49680 |
| acagattgoc | ttctctcagc | acagcaattt | tatggacctg | gtacagttct | tctgtacttt | 49740 |
| cttcaggtaa | ttttgcttat | acagactctc | tctgttcttt | tgatttaggc | acatagatcca | 49800 |
| ttaatagagt | atcacgtctc | aaaagataaa | aagaataatc | ctgtgtttcc | accaagatt | 49860 |
| aaatttactt | cataatgtgt | cctgaaaagt | tagatgtcag | ttacagtta | ataaaactcaa | 49920 |
| tctagcatga | caaatgtaaa | acgttttggt | tgtatgtttc | ggcttcagc | ggggggcttt | 49980 |
| gcatgtatcc | tgtattaatt | tgtcatgcat | cctatagata | atttggtctc | tcatatgggg | 50040 |
| agaattttgga | ctaaaacaa | attttaatt | taattggttac | tgatgtgagc | tacacatgat | 50100 |
| cttatgcttg | ttcaagagat | gattctgggt | attatccaaa | gttttaagg | tgaagtgat | 50160 |
| tttttagtgc | cctgggattt | tcttttattt | ccatagtttt | tttcgtagtc | taattttatc | 50220 |
| tttatgaaaa | tactatagac | aaataattga | aaaagacaaa | aaacgtgctg | cagtttttaa | 50280 |
| aacagaaagc | agcaatagcc | tgccttgcgc | cttccccacc | ccaagtcctc | caatttaaat | 50340 |
| tgttttgttt | tttactcttc | tgtgtttaca | cttggttttt | cagtttaagt | ctctttgat | 50400 |
| gtctactgt | agaagatgag | gatataactg | gttacctggc | ctccccagc | atctccccct | 50460 |
| acaccacaaa | agcatacaaa | agccccatct | ctttcattta | ataccattgt | atcataatgc | 50520 |
| atggctaaat | cagtgatccc | tatctatggt | atggccataa | aaactcttatt | ccagctgtca | 50580 |
| ttgaacagtg | attacactac | cacatttttt | gggtagaagg | atcattcttt | gctatctca | 50640 |
| aatgtagtat | tggccatggt | tggccttttt | aattgtttga | tttttctatg | tgcctacccc | 50700 |
| tgattttatca | ctatactctt | gctgatgtac | ctttagaaca | tgtaaaaaata | aaataaaaaa | 50760 |
| agatctcatt | gtactgtgac | agaactgaga | aatctcacaa | gtgtcacaca | cagggaagggt | 50820 |
| ctctcttggt | tgaattttct | ctcggagtgt | cccaccctag | aaacttccca | cctcctgccc | 50880 |
| ccaagtgcac | tgtttctctga | ggcctgtgct | acaactgtca | tttggaaaat | tcttttgctt | 50940 |
| tctctctata | ttgaagtctc | tgtttcccag | atccccatgc | ttcaccttcc | tatccattcc | 51000 |
| ctctttttac | tggagcaaat | tctctagatt | tatgagaaaa | gatgcctggg | agataaacttt | 51060 |
| ctctgtagct | ctcctgactg | acaaatatatt | tattttatcc | tacaatttga | ttgaattttt | 51120 |
| gggggcagg | gctatggctc | acacctgtat | ttctagcact | ttgggagacc | aaagtgggtg | 51180 |
| ggctcactga | gctcaggagt | tgcgagaccg | cctggacaac | atgggcaaac | cttgtctcta | 51240 |
| caaaaaatc | aaaaatttag | tgggtgaggt | ggcgcaagcc | tgtagtccca | gctaactcagg | 51300 |
| aggctgaggc | aggagaatgg | cttgagccta | ggaggcagag | gttgacgtga | gcggagattg | 51360 |
| caccactgca | ctccatcctg | gggtgacagag | tgagaccctg | tctcaaaaaa | aaaaaaaaaa | 51420 |
| aaagtctgac | tacagaattc | taggttcaaa | atcattttcc | attagaattt | agaagacgtt | 51480 |
| atactatttt | ttctagtgtc | cagtataata | ttgagaaggc | tagtgctggt | ataattccca | 51540 |
| ctctttgtat | gggatccggt | ttctctctta | aaaagaaaag | tatattttaag | tgtaatataa | 51600 |
| agttacagaa | acatttgcaa | atcctatgtg | tgtagctcag | tgaattatca | caaatggaac | 51660 |
| actctagag | tattgataat | attatctctt | gttcacccag | aaacactcag | gcttcttatt | 51720 |
| tctctctgt | cagtttcggc | aagttgtata | tagttttctt | ctaatactct | acttccatct | 51780 |
| gttccattct | cttgttcaat | gagaacaaga | gttaatatata | ccaataactct | agaagttcct | 51840 |
| cttgtgtttc | ccctcactca | caacctcctg | cctcttcccc | aaaggtacca | tttaccctgac | 51900 |
| ttctcaaac | atacattgca | tttgtctatt | ttgaaatgca | tatgaaattat | ttcttttttt | 51960 |
| tttttttttt | tgaaaactta | gtctctctct | gtcaccacag | ccagagtgtca | gtgttgtgat | 52020 |
| cttgggctcac | tgcgaagcaac | cgccgcctcc | cgggttcaag | caattctcct | gcctcagcct | 52080 |
| cccaagtagc | tgggattaca | ggcacgtgcc | accacgcagc | gtcaattttt | gtatttttag | 52140 |
| cagagacggg | gtttcacacg | gttgccacag | ctggctctga | actcctgacc | tcaagtgtac | 52200 |
| caccgcctc | tgcctcccaa | agtgctggga | ttacaagcat | gaaccacccg | gctcagtcac | 52260 |
| aatgtatata | aatgaaatca | caacataaat | tttaagatga | tttgggtgat | gaagaaaaata | 52320 |
| gtgaaaaact | gtgggttgca | tgtaaagtgt | tgtctcgagg | ttatagctgt | aaactctgtg | 52380 |
| ctctccacct | tagctgcacc | tagaatcact | tgcggagcct | ttgagtgtca | agccactctt | 52440 |
| cagaccattt | gactcagaac | ctctaggggt | gcaccccaaa | catttgaaaa | ctctccaggt | 52500 |
| gattccaatg | tgcagctaa | gttgagaatc | actgctttaa | aagcacatat | tagaaaaaga | 52560 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| aaaacattcc | tctcaataac | tcagaactgc | taattaaacc | taaagaagta | gaagaaagaa | 52620 |
| aaatatataat | aagagcaaaa | attaatgatc | taaggggaaga | agtacagtag | agaggatcctt | 52680 |
| aaaaactgtc | agttggatct | ttgaaaagag | taagaataat | attgacaaac | cctcggaag | 52740 |
| attaaccaag | gaaaaaagag | agaaggtata | tggatatatgt | agccaatgtc | agaaataaaa | 52800 |
| aagggggcat | tctctccaat | cccgagacg | tttttttct | ccagaagggt | atttgatgaa | 52860 |
| caactcttatg | tgaatgaatt | tgaaaaatgaa | acagatggaa | tgggaccatt | aaaaaagaaa | 52920 |
| tatgcaactt | gccaaaactg | acacagagca | agtaaggagc | ctgaatagtt | ctgtatccat | 52980 |
| ttttaaaaaca | acactaattt | aagttttgtc | atagtgaaaa | ctctaggccc | caggtggctt | 53040 |
| cattggtaac | ataacctttt | tagggccgggt | gcagtggctc | acgcctgtaa | tcccagcact | 53100 |
| ttggggaggcc | gaggtgggtg | gatcacgagg | tcaggagatc | gagaccatcc | tggctaacac | 53160 |
| ggtgaaaccc | catctctact | aaaaatacaa | aaaattagcc | aggcgcgggtg | gtggggcgtc | 53220 |
| ataatcccag | ctaactggga | ggctgagaca | ggagaatggc | gtgaacctgg | gaggcagagc | 53280 |
| ttgctgtgag | ccgagatcgc | gccactgcac | tccaaacctgg | gcaacagtg | gagactccat | 53340 |
| ctcaaaaaac | aaaaaggaga | atactgtgat | catcttaatc | aatggaggag | ataaagcatt | 53400 |
| tgtataaaat | taatacccat | tcatttttaa | aaaaattttc | tagaaaaacta | gaagagaaa | 53460 |
| cttctggaa | actgataatg | gaacaacata | aaaagtcaac | agcaaacacc | tttcttaatg | 53520 |
| gtagaatggt | agaagtttga | gtttgggaa | aagacagaga | tccactttcc | acagtttcaa | 53580 |
| ttactcaggg | tcaacctggg | tccaaaacaa | ttaagtggaa | aattccagaa | ataaacattt | 53640 |
| tataagtcct | aaattgtact | ccattctgag | tagtgtgatg | aaatgtcgtg | ccatcctgct | 53700 |
| tcgtcccacc | tggtgatagga | atcatccctt | ggtccagtgt | atccacgccc | ctatgctacc | 53760 |
| caccocactag | tgtcttggtt | atcagatcta | aaaatcatag | tatatgtagg | gttcagtaact | 53820 |
| atocattggtt | tccagcatcc | actgggaggc | tgggaatgta | acccccaga | ttaaagcaga | 53880 |
| acotatatgt | ccaaatccaa | aggatacttt | ttggttctcc | tcttcttgac | ttctcgggtg | 53940 |
| tatacaaccaa | ctgactgctt | catgagcgag | actcttccag | cogtctctga | cccatgctgc | 54000 |
| cctactcgtg | cttcttagcc | tgcctattat | gcaatctttt | ctattcagcc | tctcatggtt | 54060 |
| gattttttcc | cagaccttgc | cctggccgtt | cttctcactc | tgtgttctct | gactggactc | 54120 |
| tgcttccccc | atctgcagca | cgacatgccc | agaagtgacc | aattccatcc | ttcttaccgc | 54180 |
| gtgttcccca | cctcactaac | tgggaacatg | gctctgctgg | tccaaagcca | ttctctctgc | 54240 |
| tttactccag | caggcactcc | cttaccaggc | cccagagctg | tgcccgctcag | tgactctccc | 54300 |
| acctccactc | cccacccca | caaccctagc | atcacagtct | cgcttggacc | ttttcactgt | 54360 |
| actttggccc | cttcagaatc | cacaaaaagt | ccagagaaga | cctttaaaga | tcatgtatct | 54420 |
| agccoccttac | cttctgtgtg | tatctaaaat | aaaattcagc | gtcctgggtg | aactggcgctc | 54480 |
| cgcttttcca | tattcagctt | gcactcttct | cttccacatt | tctgtcacc | ggaccatctc | 54540 |
| tactgtcttc | aaaagcacaa | gctctttccc | tacactgagc | atttgtacat | gcttatccca | 54600 |
| ttttctggaa | cactcttctc | actacccttc | attctgtata | ctttcatacc | tttcttgatt | 54660 |
| tttccagaca | acaaacaaat | ctccaatctc | ctgacaccaa | ctcagtgctc | cacagttccag | 54720 |
| ttccattcag | acactaccca | tagttggagc | aaatcctgca | ggttaggggc | tcagtcccac | 54780 |
| aagactgctc | ccacttcaga | cactaagcac | aaatgggtcc | ccaggctacc | tcactgctac | 54840 |
| tggggccaact | acaaattccag | ggcttctggt | gactgctcac | tccaaagattt | gacaatttgg | 54900 |
| tagaacaagct | cacagaaactc | aggagagcac | tatacgtaga | attacagttt | tattatcacg | 54960 |
| ggtacaactc | aggaaacgct | aaatggaaga | catgaatagg | acaaggcatg | gggggtgggtg | 55020 |
| agcacaagctc | tttcatgccc | tctctggaca | agcctcctcc | ctgcacgtgg | atatattcac | 55080 |
| tatcctggaa | gctcccccaa | gccctatcat | tctgagtttt | tttatgtagg | tttcatata | 55140 |
| taggcgatgat | ggatgaaatc | actttctatt | ggtgactgaa | ctcatctcca | gcccccttcc | 55200 |
| ccttcttgga | ggttggggga | tggagctgaa | agttctaacc | ctctaaccac | ctagttgggt | 55260 |
| tttctgggtg | ccagcccccg | catctgaggc | tctttaggcc | cctctcccca | tgaagcattc | 55320 |
| attagcacac | aaaagacact | cctgtccctc | tgaaaaatcc | aaggggtga | ctaaattgtc | 55380 |
| tcoccttggt | ctcactcctc | gtagaactgt | attatttttc | atcgtggtag | tcatcacaat | 55440 |
| ttgaagaata | tattcatatg | tttacttget | taatatctgg | gtcggtagcc | caagcttcag | 55500 |
| gggagcagtg | cactgtctct | cggtcactcc | tatacacaca | cagctctcag | atctatgata | 55560 |
| taacttagta | agataattgt | gggaagttaa | cttagactct | tcagggtga | tcgagagtga | 55620 |
| ggaggaacca | taggcaccaa | cctctcaggg | aaaaaagctt | gtgtcaatgg | cttgccttta | 55680 |
| agggatgtag | gggtgtgatt | tgagctccag | tagacatcct | taaatctctg | gagcatgtgt | 55740 |
| atgggaataag | atggccctgc | gtacagctgc | gtacagcccg | gtgaggaccg | gccactctta | 55800 |
| ctggacatcc | agtcctctgg | cgacccaatg | acactatcag | aataagtcga | ggaagtaact | 55860 |
| ttcaggccca | tggcctcttg | tgtatgaaac | aaggacagat | ggcattttca | tcaactctcc | 55920 |
| cttgtggggg | ttgaacttct | ggaaaagaat | ccggaaactt | ggtttttcta | gacgatgttt | 55980 |
| actaactagc | gatgataaac | tttttagctaa | gaatagagtt | actttcatgg | catgggggaag | 56040 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| aaaatatattt | cccaggacaa | gttaaccatg | tctaaaagaa | tttatattagc | tttttttttt | 56100 |
| tttttttttt | ttttttttgg | ataggggtctc | tggtgccccag | gctggaacac | agtggtgacg | 56160 |
| tcatagcccca | ctgcagcctc | aaactcctga | gctggtctca | agcagtcctc | ctcacctcag | 56220 |
| tctcccaagt | agagctggga | ctgcagggcat | gcaccatcat | gccagctaa | tattttcttt | 56280 |
| ttttagagagt | tggtttctca | ctatatgtgc | caggctggtc | ttaaacttct | ggcctcaagc | 56340 |
| agtcctccca | tctcagcctc | ccaaggtgct | gggattacca | cggtgagcca | ctgctcccag | 56400 |
| cctgggtccaa | aagaatttaa | aatggagcct | ataaggaaaa | ggggaaggt | atccagataa | 56460 |
| tcttaagcac | tctgaagaat | ggctcatata | atatatggaa | aaaagaacaa | aagtcggggg | 56520 |
| gaaattcgat | gattcaaatg | agaaaatcaa | tgtaaaagtga | gttagaaaaa | atatttccag | 56580 |
| actccgggca | cagaaaagta | aactctaggt | cctgacacag | tctgttgaga | ctgtgtccta | 56640 |
| atgagatgct | tttccagcct | cagggaaggag | cctccttaaa | gagggaaata | agagactctc | 56700 |
| ctctgcttgc | ttttctgtag | cttcttgtag | ctggagcagc | aagtagaagg | gactggggaa | 56760 |
| ctgggtgtgt | caacagatct | tcactgaata | cagaagggcc | ctctcacctt | cctgaggcca | 56820 |
| tagagggaag | caggcacatc | ctgtattacg | ggtggaccaca | agagaggttt | ctgagtga | 56880 |
| ccaggggtgc | cgcctctcag | cagcagggc | aggctcttga | ccaacttaga | agctgttgcc | 56940 |
| ggtggcttca | tgcctcttct | tgtggatttg | ctgctcctt | atgtttatct | attagaataa | 57000 |
| aaataacaaa | atacaggatt | atgctgtgaa | gccaaagaaag | actgagaagg | ttattctcag | 57060 |
| acctctttac | tggtcttctc | tgttctctct | aacttctcat | tgttaggggtg | gccccaggac | 57120 |
| attctcatct | atctgtattc | actctcttgg | tgaatgtggc | caggtttatt | gcttgaataa | 57180 |
| ccatatatat | ccttatgcct | tccaaatgcc | atgcccagcc | cagacctggc | agccagcctc | 57240 |
| caggttagaa | tatctgccca | ggacatctcc | cctgggctgt | ccggttagata | tctcagggtt | 57300 |
| ctgcatgtct | aaaatggatc | tactgatattg | tccagattcc | aaagtgcctt | tcacacctct | 57360 |
| gcttccacag | ctctggtttg | agccaccatc | atctcaccta | gactgctggt | ctccctgata | 57420 |
| atgccattgc | acctctaatg | gttttctcag | caaacgttaa | atcaccaaac | tctgtttaaa | 57480 |
| aaactgcact | gggtcctttc | caatgccctc | ataatgcttg | cagtgcccca | catgatctgc | 57540 |
| ccttcacccct | caccctttct | tctctggcct | taacaccccc | actcccaaca | 57600 | |
| tgcctggcctc | ctgaggtctc | ccaagccatc | ccaggcacgg | ccttgccctg | ggctttgcac | 57660 |
| tgactctgctc | tttttccaga | cgcacttttg | caggcgctca | tgagccacct | cccatacctt | 57720 |
| ctccaaagtct | ttgttcaaat | gtcacctttg | gagaaagaaa | ggccccacaca | ctctgacctt | 57780 |
| cctgctgtgg | agtgccactg | gtactcctga | accoccttgc | cttagcatct | ttttatttct | 57840 |
| gtagcaccta | cctacctttt | ggcatgtctc | gtaatttgct | tattttactta | ctgtttgtct | 57900 |
| ccacctgtgt | gaatgttaaat | tccatgggac | agtaactctt | gcctgtttta | tctagtgttt | 57960 |
| ccatgagcct | agacaggagc | ctgacccaat | taaagtattg | attaaaaaga | caatgtaggc | 58020 |
| cagtcacgtg | agccctgtaa | tctttgccaa | ggcaggcaga | tcacttgagg | taaggagttt | 58080 |
| gagaccagac | tggtccaaagt | ggtgaaacct | cgtctccact | aaaaatacaa | aaattagctg | 58140 |
| gggtgtatgg | cgcacacctg | tagtccaagc | tactcaggag | gctaaaaagag | gagaatagaa | 58200 |
| tgccttgaa | ctgggaggca | gaggttgca | tgagctgaga | tgcgtccact | gcactccagc | 58260 |
| ctgggtgaca | gagcagactc | gtgtctcaaa | aaaaagacaa | tgtagactgt | cagaagagtt | 58320 |
| ggcaggtgct | tgatagcaga | gaggcccaga | gatgaaggct | ttccaccaag | tgacgatcct | 58380 |
| ggggggggttg | gaaggcaatc | ctattgaatt | gcctgtttgc | aggggccacca | ttggaggagt | 58440 |
| ctgcaccctt | gtacccctta | gggaggaggc | agttagcctt | gagataaggg | cagtttcagt | 58500 |
| atctatgcat | cttagagctc | ctctctccag | ccagcccagg | agggctgtgt | gggtgtcagc | 58560 |
| agttgggtgc | actgaggtgc | cattctatga | gctgaccoca | ctctacaaac | tctgagactca | 58620 |
| gccagcaaga | gagtaaccttg | atgacacttg | cagagagcct | tatgtctaca | tctagcactc | 58680 |
| tggaaaaata | ccctgtagct | agccctttt | acaagttaat | gagtaaaaaat | gcagttaatag | 58740 |
| cagttatctc | tttttttact | tgttaaagctt | attcttctct | ttaaaccgcaa | gtcagctgtt | 58800 |
| ctctttgccca | tcaactgtagc | ttttcaaato | attccacacc | tggtttatac | atgagagagc | 58860 |
| attgaaatcc | tctgcaacta | ttaaaggccac | tgggagttat | tttctcttaa | tatttttaag | 58920 |
| aaagtgtgtg | gcttgggttt | gatcatttca | aatgtgtttt | attttactta | aaaatatgta | 58980 |
| tttgtaaaaa | tattgaatat | acactaataa | ttacacacat | ttcttctgtt | tgccctctcc | 59040 |
| ttatgaaaa | ttgatgttta | cttttaataat | taactttttc | accttcttag | attttaccag | 59100 |
| taaatatctg | ctaacttgaa | taaacagcat | gactaatatt | tgtttataatt | atcttagaat | 59160 |
| tatttttctt | attttgggata | agccactgtg | ggccattctg | atgaaggcca | agtggaaggt | 59220 |
| cagggtctct | tctgaataacc | actctgggaac | gatacagctc | agaggcctgg | ttaggaagcc | 59280 |
| tggcttaaccg | aagttgtctc | actcaatacc | tcatgagccg | gcacatctgt | gtgacaaagc | 59340 |
| atcagaagg | agcccgagg | tggcagtgtg | gagctgtttt | gctgtcgtgg | ccctgagggtg | 59400 |
| tgggcaaccg | tgggagctga | ttctgggggc | tgggataata | actcagttact | ttcctttctc | 59460 |
| acagttgttt | cctctctttg | ctcctgggtg | ctgctgtggt | ttggaagatc | aaacaaagtt | 59520 |

| | | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|------------|-------|
| gttggggcctc | cagacgtaga | gaggtaagct | tcagtgggta | aagattaaag | aatccctgga | 59580 | |
| agagcttttt | ttcctctctt | ttctcttaag | caagtgggtt | ttagctattt | agtataatg | 59640 | |
| gacagacagaa | catctccacg | gaaggaaata | tcagccctt | ttctgcagctc | atgggtctgc | 59700 | |
| agcatatgggt | cagttagactt | agatgagctt | tacataaagt | gtaattagtt | ctgcaccaca | 59760 | |
| gtgtcttacc | ttccgaaaag | gaaactctgg | tattgacttg | gaaatgccag | ttttatcctt | 59820 | |
| cgacccagaaa | cccaagctcct | gactgaggcc | agagatccaa | gggcgtatta | gcacagccac | 59880 | |
| cgcggtgatta | ggcacccgct | tcatgagaac | ttgtcatgcg | agagggaaat | cagccattta | 59940 | |
| agcatctcaa | aaatttttca | ttattcaagg | aaagataaat | gtgtgtgtag | ttaagttttt | 60000 | |
| aacttgagtt | ttatttttaa | ataagtcttg | atttgtttcc | agagcttcat | tccaaaagta | 60060 | |
| ataagcaaat | atttacagct | gacaaattta | aataaaataa | cttgggttgc | aattacaaca | 60120 | |
| tatcaaatgc | ccatcaatca | atgagtggtat | aaagaaaatg | tgggggtgtg | atgtgcgtgt | 60180 | |
| gtatgtatat | atatatat | atatacat | gcagcaacct | ggatggaaat | ggagacogtt | 60240 | |
| attctaagtg | aaataattca | ggaatggaaa | accacacttc | atatattctc | acttacaagt | 60300 | |
| gggaactaag | ccctgaggat | gcaaaagcat | aagaatgata | cagtggattt | tgaggactcg | 60360 | |
| ggggaagag | tgggaggggg | atgagggata | aaatactata | catctgggtac | agtgctacat | 60420 | |
| gctcgggtga | taggtgcacc | agaatctcag | aaatcacct | gaaagaactt | attcatgtaa | 60480 | |
| ccaaacacca | cctgttcccc | caaaaacctt | tgaaaataat | aaataaataa | ataataaat | 60540 | |
| aaataaataa | atattttaaa | ataaacaccc | attaaaataa | ttacaacata | aaaaacttgt | 60600 | |
| gtattgcaac | tttcaacttc | atcatctact | tgcctttcat | ttagtcatgc | ctgtggcatt | 60660 | |
| gttggcattg | ttgtctaatt | aaatgctcac | gtgtatataa | gatctagaca | aagaagctgt | 60720 | |
| cttgggtttt | tttggatcta | ttctgcctt | ttctatata | aaaatgtatt | taattacaag | 60780 | |
| atttagaaca | gaagggaagt | atgtaaaact | acaactcaca | agcactcagc | actgttgata | 60840 | |
| ggatatctat | gtctgaatag | ggttaagaca | ggctccagga | tatgggcctc | ttgttgtggg | 60900 | |
| ccacagtacc | accttttctg | accacataaa | agatgatgtt | atcatagagg | gaatgttgca | 60960 | |
| cactctttat | tattattttt | taaatgaagt | gataacatat | tcaaaaggtt | ttttagataa | 61020 | |
| ctgtattact | gatatacttt | aaaagcttaa | tgaattacgt | gatgttctga | aaaaatttat | 61080 | |
| attttcaatt | taatttaa | gtatat | ttat | tcacaatgga | gtaagaaaga | agtaagaagc | 61140 |
| cacagtaaac | cagttataca | gaaaggagtt | gaaaactttt | gagtgctact | ctctccaaga | 61200 | |
| aggacattag | acccagggag | tttcaaggtt | tttcagagct | tatggagagt | atgatagcct | 61260 | |
| gttcctccct | gagtgcttat | gagaaataat | tgtttctaaa | atctgtaaa | caaaaggaca | 61320 | |
| gatggaaagc | ttgaaagagt | ccgaattgat | tttatacaca | tgaatcttgg | tcccaaaaacc | 61380 | |
| tgcacaaaggt | agcaccaaaa | tacaaaaggc | agctgcatag | tgaaatgtat | tgcagcagtc | 61440 | |
| taattgcagt | gtaattgaaa | agtcacttaa | taccagaaaa | cctattagaa | tatatcatgt | 61500 | |
| caaaaaaaga | atatatcatg | tcaattggga | aaaaataaca | cctgaattct | ctctagtggc | 61560 | |
| ccagaaatca | tttcgtcatt | tattgaaact | caatatccat | tcctaatttg | aaaacaaaaa | 61620 | |
| tcacagtaaa | atcaaaaata | cctttttaac | gctatgaaga | atatctggat | aaccagcatc | 61680 | |
| actcatgata | aaacaaacaa | acaccagagg | ttgtatccag | gctagttagt | tttaacaaga | 61740 | |
| gctgacatct | gagcatttaa | aaattccagg | cactgagttc | taagtgggtt | tacgtagatt | 61800 | |
| aactaatcct | tatctcaaac | ctatgaaata | ggtattgtcc | ctcgttatgt | tttatggatg | 61860 | |
| aggggactga | ggctcagaaa | ggtgaagtga | cttacaccac | ataggttaact | aagaggcaga | 61920 | |
| cgaggcctta | actagggcag | tctgtcttca | cagcccaagt | tctgttctct | cgtgtaatcc | 61980 | |
| tacagtcat | cagctttgtg | ctagatgttc | tagactgtgc | actaagtga | agggcactaa | 62040 | |
| gaagtgttgt | tactatttga | aaggaaacaa | aattattatt | ttcatataaa | gtaattaaaa | 62100 | |
| caatgtagta | gagatttgcc | tgcataaaaa | tgaaggctta | agtatatcaa | aaaataaatg | 62160 | |
| aaatacacag | aagctaaaca | atcaacttga | aattgtttgc | agtgctgcat | agagaaagta | 62220 | |
| agatgcacag | cctacaggaa | aaagattaaa | atatatatat | gtgttttagta | tggagttgggt | 62280 | |
| catctaataa | actatgcttt | taccogtgtg | atttggaatt | ttctgttaac | tgtgaagctg | 62340 | |
| accattcaga | atatatttca | tcatttttaa | aatgtcagcc | atttattatt | ccctctcaga | 62400 | |
| attgtagaag | caatgatattg | atgttatttt | taaaaaaat | agtagtagcg | gctgacgctg | 62460 | |
| tttagtccct | gctacacacc | agccccggtt | ttaatgtgct | ttaacacatg | tgccctctga | 62520 | |
| acccctgtctt | attgtgaccc | cagcacttaa | gcacagaccc | taaggagcat | tgtgcatatt | 62580 | |
| attagtcgag | taaatggacc | tacagcttta | aaaaatgtaa | gaatgttggt | attttaagaa | 62640 | |
| aatatttata | aaaatgcctt | ctagaacag | agccatatct | atccaggcca | agattctctc | 62700 | |
| ttctacagta | gttccacagg | ggcagtttga | agtaaggcct | acttgtatcc | caaggcataa | 62760 | |
| cctcaaaagt | tcagcatcat | gcacagaaag | atctctaata | tcccttaatga | tatgtgaacg | 62820 | |
| gctgttccac | atgtgctagg | cctgtgtgctg | tagtatctca | tttgatacca | gtgacaacct | 62880 | |
| tatgggctag | gaacacattt | cagaggagg | acacggatgc | ccagagactt | caaacacctt | 62940 | |
| tcttaaggtc | acacagctgt | agactatttg | agcaggagtt | cacatcaagc | ctccctaaat | 63000 | |

| | | | | | | |
|--------------|-------------|-------------|-------------|-------------|-------------|-------|
| ccccgagcccc | tgttttgtccc | tctactctgc | aaaatttcac | atttgtttgt | ttttagctta | 63060 |
| taccacctct | ctgaaggaag | gggtctttca | gttccatctc | gccgtataat | gaggcatagg | 63120 |
| tttttgttgt | cctgtgaatc | cttgggtgcca | gtttcagtag | gaagaaagca | agcctgaact | 63180 |
| tagtagtaga | aatacggaga | gataggggcca | caccatttta | tgtatttgac | aagtcgacgt | 63240 |
| catagcatta | aatccacagt | accttttattt | ttatctttat | tttcttattt | tactttaagt | 63300 |
| tctgggatac | atgtacagaa | catgtagatt | tgttacatag | gtatacatgt | gccgtgggtg | 63360 |
| tttgcgcac | ctatccagct | gtttacttag | ttttaagccc | tgcattgcatt | tggattttgt | 63420 |
| cctaattgctc | tcctctccctc | tgctcccccac | ccctgcacagg | ccccagtggtg | tgatgttccc | 63480 |
| tctcctgtgt | ccatgtgttc | actcggttca | actcccactt | aagagtggagg | acatggagta | 63540 |
| tttgggtttc | tgttcctgtg | ttgggtttgct | gagactgatg | gcttccagct | tcactcgtgt | 63600 |
| ccctgcacaa | gatattgaact | cattttatttt | atggcccgcat | agtattccat | gggtgatattg | 63660 |
| tgcctagtgtt | cttttatcca | gtctatcatt | gatgggcatt | tgggttggtt | ctaagtcttc | 63720 |
| gctactgtaa | atagtgtctgc | aataaacatc | gtacgttaat | atgtatctct | gacaggggct | 63780 |
| tttaaaaaat | ataaggacaa | tggtgtttatc | acatccaaca | aaatttagta | atccctaatg | 63840 |
| ttatctaata | tcaagtctcaa | aaagtatcat | tttggacttg | atttgtttga | atcaggagg | 63900 |
| atgagggtgtt | tttaacatgt | agcatttaact | aataaaaacc | cttcagtact | catttccatta | 63960 |
| atggccttgc | catattatttt | atcagcaact | tcttcgagag | atgcaacaga | tggccagccg | 64020 |
| tcctcttgcc | tctgttaattg | tcgccttgga | aacagatgag | gagcctcctg | atcttattgg | 64080 |
| ggggagatata | aagggtgagaa | tgtgtactcag | aagtccctat | aacttgaact | tttaaaaactt | 64140 |
| aggtcctcaa | gtctgggaaa | ccagagagag | caaaagccct | attccattac | ctcctttttt | 64200 |
| tcctctcttt | ctttttaaca | ataagctaaa | acctgaagtt | gctgggtact | cttttgcctt | 64260 |
| tgtttttcaa | tcttttgttct | atgtcgctga | gcagtcctgt | gttcaggggga | ttttaaaatc | 64320 |
| taaatgatttt | gtctgctttt | cttttaagta | ttctgtggca | tacaaaaact | ttatttgaaa | 64380 |
| agagctacat | tgaacatcga | atcatttgtt | ttttgtttgt | ttgtttgttt | gagacagagt | 64440 |
| cttactgtgt | cacacaggtc | ggagccagct | ggcatgatct | cagctcactg | caactcactc | 64500 |
| cccggtgttca | agcgattctc | ttgcgccagc | ctcctgagtt | gctgggatta | ccgattgtgcg | 64560 |
| caaccaacgccc | cagctgattt | ttgtattttt | agtgaggatg | gggtttcacc | atgttgccca | 64620 |
| ggctgggtctc | gaactcctga | cttcaagcaa | tcacccacc | tcggcctccc | aaagtgcctg | 64680 |
| gattcacaggt | gtgaaccacc | gcaccagccc | tatttttctt | tatatttttg | ggaatttttt | 64740 |
| aagcagcaaaa | atgataatac | actgtgtatt | aatactctgt | ggaataatatt | tatcaccccc | 64800 |
| ctccccaaaa | tgatgagaac | aatgtgaaga | ctttctcctt | gggcaaaatg | atggaatttt | 64860 |
| taagatacgc | tggagaaata | gctatgtatc | ttgaataaaa | catacttact | taaaatgtta | 64920 |
| ctagagctcac | aggaagtagg | taaaatctcc | aggggaatac | ctcctctccc | accctaaaaa | 64980 |
| gaacaaaccca | taaaccacaa | ccagaacctat | gatttgtcct | aagtgtgttc | agatgttaaa | 65040 |
| cacaacgccc | acctgaatag | atggaaatcc | tctctataga | aaaaatctct | caattcagcg | 65100 |
| ctgttggaagt | cccacagatc | aagatgaagt | aattaaagat | caccagggac | aagaggcaag | 65160 |
| ccgccataaaa | taagagtgc | agaaacaaca | gatgattgat | gtggacagga | ctccaaatgt | 65220 |
| tagaattgaa | atctaaagaa | tataatataa | ctgtgtatga | aatgtttaaa | gaaataaaag | 65280 |
| atgaaatcat | aaagatgaga | gtcaagtatg | gccaaagcaga | tctgaagaag | aaccaaattg | 65340 |
| aactcttaga | aatgaaaaat | atacttgttg | aaatttttta | aataaataaa | ctcaattttac | 65400 |
| atgtttaaaca | gcatatcaga | cataagagaa | tttacaaaagt | ggaagatgaa | tctgaaaaaa | 65460 |
| gtacgcagaaa | tgcagctcag | agagacaagg | agatagaaaag | tacaatagat | actaagaata | 65520 |
| tggaggatag | aatgagaaga | tccaactcat | atatatttcc | agaatcccag | aacaaaaggag | 65580 |
| aggaatattt | cagaaatgat | agctttccag | aactgatgga | aaacatgaat | atacagattc | 65640 |
| cagaagcaaaa | acatattgta | agcatgatta | aagaaagaaa | gggaaggcag | gaaggaaagc | 65700 |
| aaacaagtgtg | gatttctcat | aatgaatttg | tataactcca | aagcaaaaga | gaacatctta | 65760 |
| aaatcagcca | gagaaaaagt | atcacataca | acgtggagaa | actgaacggt | agtgcaactgt | 65820 |
| tgttggaaagt | ataaagtgat | ggagccactg | tggaaaaacag | taccactatt | ccccccagaa | 65880 |
| ttaaaagttag | aattatatac | gatctggcaa | tctgggttta | tacatagaag | aactgaaagc | 65940 |
| agagtctcaa | agacataatt | gaacatcagt | gttcttggca | gcattattca | cagtagccaa | 66000 |
| aaagtggaa | caggcccaag | tgtccatgga | tggataaaatg | gataagcaca | acgtgtgtcta | 66060 |
| tgcattccaat | ggaattgtat | ccagocctcc | aaaggaagca | aattctgcac | catgtctaca | 66120 |
| tgtgatgaac | cttgaagaca | tactaagtga | aatactccag | taacaaaaaa | ataaacattg | 66180 |
| tatgattaca | cttttaagag | atacacttgt | ggtagtcaag | ctcataaaga | caggaagttag | 66240 |
| aaacgtgggt | cacaggggct | ggggaagagg | gaaaatgggg | agttagtgtt | taatgggtac | 66300 |
| gaagtttgag | ttttacaaga | tctggagatg | ctgtgagatg | atgtgtaagt | atggttgcac | 66360 |
| aaacatgtga | atatacttaa | taccactgaa | ctgtacattt | ttaaattggtc | aagatggtaa | 66420 |
| atgttatgtg | tattttatca | caatttttgt | taaaaaatggg | aaaaggagcc | gggtgtggtg | 66480 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gctaaccacct | gtaatccgag | cactttggga | ggccgaggcg | ggcggatcac | ttgaggtcag | 66540 |
| gagttcaaga | ccagcctggc | caacatggtg | aaaccccatc | tctactaaaa | atacagaaat | 66600 |
| tagccaggga | tgatgggcaca | catctgtaat | cccagctact | tggggaggtg | aggcatgaga | 66660 |
| atcatttgaa | cctagaagcg | agagttcaca | gtgagctgag | attgcaccac | tgtaactccag | 66720 |
| cctggggcgac | agaacaagac | tctgtcgcaa | aaaaaaaaaa | aaaaaaaaaa | aaaagggaaa | 66780 |
| agatcacctg | caaaggaatg | acaagctgac | cgctaaactc | tcagaagtca | gaagagagtg | 66840 |
| agagaatgtc | tttgaagtac | tgaagaagac | ctccccacct | agaatagtgt | gcccagtaag | 66900 |
| caccttccga | gaacaaggat | gaaataaaga | tatcttcaaa | ttagaataaa | aattgagaat | 66960 |
| ttgccaccag | cagacctgca | caaaagaaat | gtctgaagga | tgtgttaaaa | tgaaggaaaa | 67020 |
| ttacccccag | gaggatagac | taaagtgtaa | gaaagaatag | tgagtaataa | aaagtataat | 67080 |
| catatataga | aatctaaaca | aacattatcc | tggtaaaata | atattttgtg | gttaaaaaat | 67140 |
| agacaagcgt | aaagtattga | acaacatgat | atatgtcagc | agcatatgac | acagttagtt | 67200 |
| gctccttaaa | acattttctt | gattttcttc | caggacacca | cacacttgcc | tctagctgtg | 67260 |
| tcttgctcag | tgggtctttg | tcagtcctct | tcatcttttg | acagttgggt | gcgtcaggcc | 67320 |
| ctgccttttg | ccccttcttg | tgtcttttac | accactttct | tcagtgatct | tcactatcac | 67380 |
| aggtttaagt | accatgatcc | tctcacctgg | actgtctcca | tgagcctcca | acacaaacat | 67440 |
| cagbtgccta | ctcagtgctc | ccacttggat | atcttaaacat | ggcccatggc | aaactgagct | 67500 |
| ttggagtgtt | atttctcaga | cttgtctctc | tgacataacct | tcogcctttc | agcacataac | 67560 |
| acctccatca | ccctagttgc | tcaggccaaa | atcgtggaat | tgttcttgac | tcattcttga | 67620 |
| tgtcatgttc | tcattgggaca | gccagttctat | cagcaaatct | tgactccgta | ttcaaaaaat | 67680 |
| tttccccagc | actgtgtccc | atctctgcca | ctattcccac | ttgatctagg | ccaccgtgtg | 67740 |
| ctctcacctg | gattcttgca | gttacttcat | aactgatctc | cctgcctctg | ccattcaccc | 67800 |
| atctagtccg | ttctcttgtc | ccagcagag | tggtactgga | aattattttc | ttcagccaaa | 67860 |
| ctgaagaaac | tgatgacatt | tgtgggtata | gtcagcagaa | ttctaaaggt | ggccctcaag | 67920 |
| atttccatcc | cacgttttatt | cagtcacaca | ctactctagg | tactgctaga | gggattttgc | 67980 |
| agatgtaatg | aaacaaggga | gttgacttta | agatacagaa | agttttctctg | gctggtggca | 68040 |
| gaaagaggaa | gccagagaga | ttggaacat | gagaaggact | tgacacacca | tccttggttt | 68100 |
| agagatggag | tggccacaga | agtttttttg | tttttgtgag | acagagctctg | gctgtgtcac | 68160 |
| ccacactgga | tgtcagtggc | atgatcatcg | ctcactgtag | ctccaaactc | ttggggctcaa | 68220 |
| gggatctctc | ctcttgggtc | tcocaaagtg | ctgggattac | agggtgtgagc | cactgcgtaa | 68280 |
| gccctgaatg | cattttcaag | cagagtcttc | cccagagtct | ccaactaaga | gtcccatcca | 68340 |
| gctgacaaat | tgacttcagc | ctgagagacc | tcaggagata | ctgtgataatt | cccagaaact | 68400 |
| catggaagac | atgaacagac | ctgtctctct | cggcagagaa | ccctacagag | ctgtctggga | 68460 |
| cttctgacct | ctaaagctgt | gagataataa | gtgggtgttg | ttttaagcaa | ctaagtttgc | 68520 |
| agtagtttgt | tacactaata | taagctgggt | gcggtaggtc | acacctttaa | tcaccgacct | 68580 |
| ttggggaggct | gagggtggag | gatcgcttga | gcccaggagt | tcaagacctg | cctggacaac | 68640 |
| atagttagac | cccattttcta | ttctttaaaa | aaaaaaagga | aaggaaagaa | aaagagaaga | 68700 |
| aactaatata | ggggacaaat | ctcatctaaa | ggctgcaaca | gaagatgaaa | acatattttc | 68760 |
| cagagccaaa | tgcattccaga | gagtagacaa | agaatctact | tgggtcagcca | accagcctt | 68820 |
| tacaagcttt | tacatccagc | aggcttgtct | acttagtgtc | taaaggacag | acaaaataga | 68880 |
| atgtttttta | agtcctccatg | atatctgcat | ctgtgtttga | gccaggttatg | cagaactatc | 68940 |
| tgacaaatct | accaagcact | caccactgcc | attctctgtga | cctacagtta | tgaattagaa | 69000 |
| tttcaaatcg | agccaaagta | cagatccag | aagaaaccca | cttggcgtgg | ccccttccct | 69060 |
| acagttcagg | cagccctctg | gaaggggact | gcgagacctc | agtggaaggt | agcaagcagt | 69120 |
| accaccgctg | tcactgtgtg | ctgttagctc | ctctgtgcc | tctctgtccc | tttttcaata | 69180 |
| tcttaaaaac | tgcaatgttac | cttttttata | tacctgaaat | tgatgggtct | tataagcaca | 69240 |
| aatgaccggt | aaagttgtcog | tatacattaa | agcttgagta | tcacttaagg | taatttttgt | 69300 |
| ttctcttgcc | agactgttcc | caaacccatt | gcactggagc | cggtgttttg | caacaaagcc | 69360 |
| gctgtcctct | ctgtgtttgt | gaggtccct | cgaggcctgg | gtggcatccc | tcctcctggg | 69420 |
| cagtcagggtg | agtagatgco | gtccagcgaa | agacaccttc | taagcatgtg | agggagctaa | 69480 |
| gcattgggata | cttccctttc | ctagagaaca | aggttataaa | ggtgataacc | aacagtgctc | 69540 |
| gcccatgccca | tggcaaggat | tcagtgagat | aaccttattt | cacctatatg | cccaaaactg | 69600 |
| ttttcccgag | gtcatcgctc | cttccctggg | ggacgtcccg | ggacctctct | accgttctg | 69660 |
| ctctccccag | ctctcagtcg | tcgtgtgtgc | ctctgaccac | acctctcttt | gcgtgtggag | 69720 |
| ttctctggtg | cccttttgct | ctctgtctct | ttccctgtct | cttttgccct | gtctctcccc | 69780 |
| tgaactctggg | tgtgtctcct | tcacctcgca | tgttcttctc | cagttgcctt | ctccactcca | 69840 |
| caacttccag | ccatctcccc | tgggctggtg | tttcttcaat | tgggtatttt | agccctcctg | 69900 |
| cattttctttt | gctttttcac | tgttttaaat | ataaggcaca | aacacaacga | taacataaag | 69960 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gaaacgtctg | aaaaatccacg | cacagttctg | ctatcccaca | accaccacat | tttcattggt | 70020 |
| ctgtgttctt | tccagccctg | gccagcaatt | tagttgaaat | gatttgtatag | atcctaatttt | 70080 |
| tattctgtctt | ttttcatata | actatcttaa | acatttttat | agcctactac | ataaatatttt | 70140 |
| accaagtcat | gtgacacttc | tccatcfaat | atatgaatca | tgattatata | aaactttctt | 70200 |
| ttgaagacag | aaaatatgga | atacagtttt | gaaagtatct | cgcttaatta | cagagtgtct | 70260 |
| tgaggatgac | ctcaatgtga | tcatttttga | caactttcag | ttgagatgat | tttggagtca | 70320 |
| gggttacctt | agccagagag | caccctgtgt | ctgggtgaaa | ggctgcccca | ggaatactggt | 70380 |
| ccaggccctg | agctccactg | tcccttctag | gcaccctcag | accocagaaga | tccctcaagcc | 70440 |
| cagcttttaga | atggtttagga | tctatagtat | cttcacgagg | gaggcccttc | gacatggctc | 70500 |
| ggccaggatt | cattgtcttc | atctcaccta | gggattccctg | tgagttgagc | ccctaactcg | 70560 |
| gccttgctg | tgttgaggac | aggacaactc | tctcgtctgt | cactcaagtg | cagttttgtg | 70620 |
| ctctctccac | cgccccagcc | agcaaacagt | ggctggtgtg | ccctgagggc | aggggtggcgg | 70680 |
| ctggggcgcg | aaggtagacg | gtgtgcagcc | ctgtagcctt | tcctgtctca | gcctttgcct | 70740 |
| gacatgcccc | aaccacctga | aacacctaat | aaaggtctgc | tcaatgcttg | agaccacccc | 70800 |
| aaacagtggg | ccgtccttgc | tgccccctgt | gtacccccctg | ttttaagttg | tgcccccttc | 70860 |
| taggctccca | gagcattccc | cactgtgtgc | caggcactta | cagctgtcta | ccccaccccc | 70920 |
| cactcccctt | cacaggctca | aacacggcaag | taacacgctc | cactaacaca | actcctggga | 70980 |
| ggctccagca | gagctccgag | ccctgcctgg | ccatttccca | cgtatatag | gaatccagct | 71040 |
| cccccagaac | cgtgccttcc | ctaattggatg | tcagattact | ccgttatcaa | tacagggaag | 71100 |
| aaaaatggat | tgcaagtccc | caaaagcaag | tgggctggcg | ggcagtgatg | agcataaagg | 71160 |
| gtagttgcta | ttgatctggtg | gttgctgtgc | gccagatact | aagactttat | gcatggccagg | 71220 |
| tactaaagact | ttgtcagttc | actcaattgat | tcttcacact | agccctgtac | tactattatc | 71280 |
| cccattttac | aaatgaggaa | actgaggtgc | agggaggtta | aagaactcct | gttattccaa | 71340 |
| agcctggccc | cagagcctgc | cctgcctgta | tcgtggtatg | gaattggaagt | gactaatgct | 71400 |
| gactgagtga | taggggtgaa | tctgaatctc | gtcataaaagc | gtgctgctat | tctcctggct | 71460 |
| cttttggagg | aaggcaggag | aaatttggca | cttagcacac | gtctacaact | tcatttatat | 71520 |
| gttcccacca | ccaagcagag | cactgtgctt | ggttgatgtg | ctgaaaagca | aactgcacct | 71580 |
| tttgtgtaaa | atgtagggaa | aggagaagaa | tgtgtatttg | tatttgcata | agbtataccct | 71640 |
| caagaataac | ataaaaattta | gtgaaagctg | ttaccagagt | ggagggggaag | gcggaggggag | 71700 |
| ggtcagggtg | gaagggaagac | ttttccctgt | gtatgtgtac | ctttttgggg | gttttttgtt | 71760 |
| ctggtttttt | ggttttgttg | ttttgcacca | tgggattgaa | aaataaatgt | ttaaattatt | 71820 |
| ttttaaaaca | aaataaaaagc | cagctgtctc | cagcatatgt | tctctctgtg | gttctcccaa | 71880 |
| ggctcttgctg | tggccagcgc | cctgggtgac | atttctcagc | agatgcagat | abgtgtacaa | 71940 |
| gagaagtcat | gagccgtgag | aaaccggaaag | cagcagcccc | ctgcacagcc | tgggacctgc | 72000 |
| atctgatgct | ggggccaggg | actctcccac | gcacgagcta | gtgagtggca | caccagagcc | 72060 |
| atctgcaggg | aagggctgtg | cggggaaatg | gctgtgcggt | gcgggaacgga | agactggaaa | 72120 |
| ccctcaaagc | atctgactca | cctgcatgat | cacaagcttt | ctttgacggt | ttctcccate | 72180 |
| cgtgttccag | catctaacct | tttacttttg | cataggaaat | acttgattta | attacaggtc | 72240 |
| cagggatgag | ctgatggttg | ctggaggagg | ccagtgtaga | gccagttaga | gaactaggaa | 72300 |
| tgacactcag | gttcaactgtg | gaaaactgtt | cttgggactg | tctcaactgt | gcaaaaaaca | 72360 |
| aaagatggag | tgttttacaag | tagacattcg | tcactcagttg | ttcttgaaac | tggtctttta | 72420 |
| aaaactagtc | agatgaatta | acttgttttc | atctgaagcc | tgctattctt | tttaaaagat | 72480 |
| gtgctattta | ttcttgcaag | atttaggcaa | ttatctctct | tcacgggagt | accttttttt | 72540 |
| ctagttgaga | attaataatg | gtccatctct | tttgatcata | tcaagctagg | atagaagggg | 72600 |
| ggctatttta | aatgtcaagg | tcagcagttg | tactttgaat | gtaaactggt | ataataggta | 72660 |
| gtttttctata | gtaacttgat | taatttagtc | ttaatccatt | tgaaactctc | tcttctcttc | 72720 |
| ttctctcctg | ttccctctct | ttccatctct | acctctccct | ttctcacacat | acacacacaa | 72780 |
| acacatacac | acacactcaa | gtgcctagac | tttaaataga | tctagcaatt | ggaaaagttag | 72840 |
| taagcctaag | tttttacata | attgcattcc | tacattcttg | taaaatttaa | atagctacca | 72900 |
| ttggcaatct | gctttttttc | taaaactctga | tttgacgcca | ggaaagaatt | ttctcaccca | 72960 |
| aggaacattt | gatctagcag | cagggtatgag | aggaagcgag | aaatgaatga | actgtgaaag | 73020 |
| ctcctgtttt | tattatcaaa | aaggacactg | tcaagaaggc | gccccctgoc | ccccaccccg | 73080 |
| tgccacccta | ggcctgataa | gcgatccagag | gaaaggactc | attcatgtca | cgcttctctg | 73140 |
| agcagaaaaa | agcactgaga | gcacttggga | ccccctggatc | agagagcatc | tgctgttctc | 73200 |
| gcagcctcct | ctgaactctg | gggtcattct | caggctgggg | tggactcaga | tgccagggaaa | 73260 |
| gggacagcct | cccatgttca | ggcagaagct | goccaaagcc | tggagaagga | ctctgtttgc | 73320 |
| ctctttcccc | caggaggggc | tcgaccacc | cacctctcct | ctcagaccaa | gggtggtggct | 73380 |
| gtgaggaggg | cagcaaatgc | tgacaaggat | gaaaagcaca | tggaaaaaaa | tggacgagga | 73440 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|------------|-------|
| gggaaaactc | tgccaaatgg | aaaatgacca | aatttaagag | ggtgggacag | tccctgctc | 73500 |
| ctctcccaga | gggcactgct | tggaattgt | gttttccca | tttatggtgc | tctgtattct | 73560 |
| ggcattatgc | agcagcctcc | cagaagctct | cttctgcttc | aaaacctggg | atctctggga | 73620 |
| ttaccctatt | gggatggacc | gttggacagc | aatgctcgag | tttgtgaatt | tgagagata | 73680 |
| ctcaaaagag | ctaaaaactgc | agcattttac | ctttaaatgc | agtgcctaga | gagagagtat | 73740 |
| tgtctcttcc | ccaacactaa | ccccactccc | atgaagaatt | gcctggaaag | atgttttcaa | 73800 |
| ggaaattgaa | ccataaaaaa | ctatctgatg | cacagaaacac | ctctactttg | agactcacct | 73860 |
| ctcataaagc | ttctttttca | cattactgtt | aaagaccaga | cgttctagaa | aagacccttc | 73920 |
| ctctcatgag | ctcccccatc | cctgctacag | aacacagcac | ccatggcgcc | tgagtgagg | 73980 |
| tggcccttta | attcccacag | gcccccccag | caaggccaaa | gggaggcccc | tggtatttgt | 74040 |
| cctcctacaa | ggaagatcct | ctttgtttgt | tcaaggacc | agttttccta | ggccaaagaa | 74100 |
| gtctctctcc | catgttagtc | ctatgccttg | aaatatcatg | caccatgacc | caagccatc | 74160 |
| tggttatgct | ttattttttt | cctaaaagat | aatgtttatt | tttaaaaagg | aaggaaggag | 74220 |
| caagtgaagt | ttcattctgc | tccagcggtg | gggaagccgc | tgaatccacc | tgtttctcct | 74280 |
| ttgcaaccag | cagcaaaacag | ctttctccg | cctcaggga | gaaaaaggga | atggcaggga | 74340 |
| gtaagaggcg | ctgggctcgg | agcctgtttc | caagaaggaa | ttggttgtca | tctggcagtg | 74400 |
| ttgcgctgca | caagagagcc | tgtatataaa | ttaaaaatgt | caagacaaca | ctgaccttgc | 74460 |
| acttgtacat | aactatacag | tagtgtccag | aatgttcaga | catctggagt | gtacataaaa | 74520 |
| cagaaaaaat | cttcatgtat | ttttattaaa | tataacaatg | tctgagtttc | acctaaagt | 74580 |
| tttttgtgcc | atatgctgga | tatccaggtt | ctcgccaggc | cccgatatac | gaataacaaa | 74640 |
| cccaagaaac | gcatccccat | tgtctagtgt | gttcagatgc | atcttggcac | aattagtgat | 74700 |
| ttctttaaac | aggactcatc | tgtcagagtg | cacatgaaaa | atcaggcgag | gtacgaaac | 74760 |
| gacagcgctg | gaggagatc | aggaagcaga | ggcgctccctg | ccgctgcctt | tggccctgca | 74820 |
| agcacatcat | gaccctttct | ggcagcctct | tgggtctctg | ggtagtggag | gatgaccagt | 74880 |
| cttctgtcct | gaaatgttct | tcttagtctt | taagttcaaa | gactaacctg | tagcaatcag | 74940 |
| actttccaaa | agggggttct | ccattttttg | tagttttgtc | taaaatttta | atgaccttat | 75000 |
| cctggaatca | gtttattata | ctgaaaactg | gggggtggag | tagggagcta | gtttgttgat | 75060 |
| aaatagttcc | catttccccg | tggagaattt | gacataacct | ggactcctgt | gtgcctcctg | 75120 |
| ccatccctgc | acacagcctg | gggagaagcc | tgtgcctccc | ctgttgggaa | gaaggcaacc | 75180 |
| ccagatcccc | tgagctaac | cggaggaaaag | gcagtcctgg | acagaagact | gtcagcagaa | 75240 |
| ggaaagtact | ggactaccgc | tgggttaagtc | ctgccattca | agactggaga | cacctgggaa | 75300 |
| ataaaaagag | cagggcactg | ctgttgggaa | gaggcathtt | acottccagt | gcaaatcctg | 75360 |
| ctcctttgat | ttaatggggg | gtactggggc | caggggctga | tccacttcct | tgggagatgg | 75420 |
| tgggtgtttc | atgaacatct | ttgatccttc | catttcattt | attcatccat | ccattcaaca | 75480 |
| agtatttgct | aaacactaac | ttaagcta | gctagggtag | tgactgagat | gtaaaaatag | 75540 |
| atttttagaat | taaaaacaaa | tccaaagtct | cacacccctg | tcatccagg | agatctttcc | 75600 |
| tttgtgtggt | ttctgtgaga | atttggccat | ctgaggacac | agccaggacg | gcagaggcct | 75660 |
| cctggcctca | gggcattgcc | tgcctacctt | ctgaaatgtt | taccccatg | accaaacttg | 75720 |
| gtccacagca | tttgggtggt | ttctagatag | ccaggcccac | caagagatat | tgccccctga | 75780 |
| tgagagtcaa | acaccctgcc | tacaaggaga | tgttttgaaa | tggagaggaa | aattggcacc | 75840 |
| tcatctttta | aaggcagtaa | tggaaattgat | tttcagtaac | tgaatttgtg | cacaaaacat | 75900 |
| tctaaacact | agtgaagcct | tttctgttga | actaattaat | tctggctctg | gaaatgtttt | 75960 |
| tgttttatag | ttattttacg | tgttctttgt | ttggattcaa | gcttagtttg | tttaatatga | 76020 |
| taattttagca | tctattacac | tcatgtaaat | atggagtaag | tatttgaaac | tatttcatgt | 76080 |
| cggggattgt | gggtgttata | catacattta | ggactgcaat | tttttgggtat | ttttttgtat | 76140 |
| tgtaaaaata | cagctaattt | aagcaggaac | agagagaacta | agggagggtc | gtgcatttta | 76200 |
| aaacacaaat | tgaagaactt | gtatataaac | aaaagtaaat | actataatac | aaacttctct | 76260 |
| ctgaataaaa | agtagactctg | gtaaaaatgt | ggcttttgtt | ctgagtgttt | ctctttgatt | 76320 |
| ttgcatttgt | ttgcctatat | ttacattttg | gtcatttagaa | ccttaaggga | aaaaaaaag | 76380 |
| cgtttaccag | ttttcagact | gcgtcagagg | gagcacttga | cagtttaacgg | aggtgagggt | 76440 |
| gaaaaacccc | ggcagggtacc | ctccagggcc | ctccagggcag | ctttagagcc | acagaccttc | 76500 |
| cacacagcct | tttgtgtcct | acagggcagc | tgggcccctg | ggggcaggga | cagatctgag | 76560 |
| cagtgaggta | tggacagccg | tctcatgct | ccactaaacca | caccactgtc | tccagtggtg | 76620 |
| ctgtttatcc | tccactttgc | gtctctctca | caggctcctg | aatgtcctct | aggtttctct | 76680 |
| caaatgtcta | aatctcaggt | gtgctgacag | gtctcagccg | tggccagtg | ttcctcggc | 76740 |
| ctgaagtgtt | tgaaggaagt | actcccattg | gcaaggccct | gtttgggggc | ctctgtgtgt | 76800 |
| tcttgccccc | ttttatttgg | ctcctgcctt | cagaatgggt | cagagtcag | tatttctagt | 76860 |
| cagagtcgag | attcccagtg | ctagcagttc | cttccagaag | gatgaggtca | ttgaaagaa | 76920 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gaggaaacac | aagacgggtca | ctcctgtagg | ggaagatggc | agggtactag | ggccaggggcc | 76980 |
| tgaggtgaag | gcaggagcct | ggaggaggcc | tgccctaccag | agctcagggtc | tggtcttaggg | 77040 |
| agggaaatgg | ctggactagt | ggctgcaaac | aagttagagaa | ggaccttgaa | actcttccag | 77100 |
| ggagggttat | gagagccacc | aacagtaaca | ggatagaaac | cctgtgattt | gttggtactg | 77160 |
| tggcagagac | aatgctgtgt | ttaccaaac | tactgggccc | acagctcatc | atttccccag | 77220 |
| ctccctgcag | ttaggttggg | accaagctaa | gccatggcca | aggaaagggt | ggaagaagtg | 77280 |
| atgggcacaa | attccaaacg | tggcacacaa | atgctcccac | aggggactct | cctaccggtc | 77340 |
| acatgtttaa | gactggaggg | actgtgggtc | cctgagtcac | cccttagagg | ccggccgctc | 77400 |
| gatcagtaag | agctgtgtcg | cctgtgtcat | cactggaaag | caaacctact | ctcgtgatcc | 77460 |
| actgagattt | ggggctttgt | tacactagtt | agcttttctt | attctcacta | atatacaaat | 77520 |
| cagcactttg | aagtgcagtg | cacatgtaac | caaaccttta | agtacatggc | actgggtcag | 77580 |
| cgtgtggcca | gggtggcagc | agtgtataga | cagggtgcagc | tggcaggaca | actgaggtcc | 77640 |
| ctgccaggcg | gagcaaaaca | tctgttaagg | ccgtttcact | cgataactcg | gaaaacaaac | 77700 |
| tccatgtctg | ctgacttttt | gacttgaggg | gaataaacaa | caacaacaac | aacaaatatg | 77760 |
| tgttgattgt | tactggctgc | ttttggcgag | gcatacagag | aaaaagttta | cttcaagaaa | 77820 |
| gaattagcca | gtttacaagc | agaagggaaa | aagatggaaa | gatgttaacc | agggccttgc | 77880 |
| aattctggaa | agagccaact | cttctagtct | ccaaccagta | agaggtgcag | gtgagcaaca | 77940 |
| aagggcccat | gagacttagt | tgaacacagt | gactggggct | catggtgacc | atccaatgaa | 78000 |
| gggtgtgtgt | ttcccactgg | agcatggaaa | ctgcaaagcc | accagcttga | gggaggggca | 78060 |
| cgtggttaag | gaagcaaaag | aaaagctgac | atgagaactg | tttccaggaa | caaacctaga | 78120 |
| atgtagtcac | tgacccttga | cctgcctgga | aataaataaa | gatgtagtaa | atttaagaaa | 78180 |
| ccattagcaa | agccttaaaa | aaaaaaaaaa | aaaaaaaaaa | aaacggtggc | aattggaggaa | 78240 |
| gtcttaaatc | caccttcagg | caggaaaatg | gttgtgaaag | ctgtgtccag | cgagggcttt | 78300 |
| acaagctcag | atgtgggccc | aagagcagaa | ctgggcttga | gcgaggactc | ccccagcccc | 78360 |
| aggacagggc | ccagcaggat | tgcaaaaaca | ccgcagagca | ccaattgtgt | catctgcacc | 78420 |
| cgctccttcc | tgggtaggag | ggcctaggag | gggtgcctgc | cgcttctcat | cactacctgc | 78480 |
| tgggtgggag | gtggattaa | ctgtcttgtt | aattcaggat | tgtggaactt | aaagagccac | 78540 |
| ctgagggagg | ctgcataccc | tggagggtct | ggcgggggtg | gacctcgat | tgtcccctca | 78600 |
| ggagaggggt | cctgtgacct | atgtaggaga | ggaagcacaa | aactgggttt | ggtgatcaga | 78660 |
| atagcaaaat | gggccagatg | cttacacac | atttctctct | cctgagtgca | cagctctcag | 78720 |
| tgcattggct | gggcccgctg | catctacca | gatgccttgg | gaaccgacct | accaatgat | 78780 |
| gtgtcatttc | aaggccattc | acagtgacag | tgtctcctcc | actgtctcca | tcaaccttga | 78840 |
| aggtctatgg | tgaagatgg | aatacaacag | gattgaagga | gaactgggtc | ctgaatgact | 78900 |
| ttgtggaaca | gagcactata | ggccaaagg | agtagaatta | ggtatcttgt | tttatgtctc | 78960 |
| ttaaagtgaat | taattaatc | tctagattta | ttcacctagt | tattcaacaa | atttttattg | 79020 |
| agcacatcag | ggctagggca | tttgttcagg | atcacagagt | ttgcaatgaa | caaaaacagaa | 79080 |
| aactccatgt | cctccagggg | agttgataat | aaaatgtcag | ataagctgt | tgaagaaagg | 79140 |
| taaaagttag | gttaggggta | gaaagtgtcg | gggggattgc | tacttaacat | agtgtggcca | 79200 |
| ggggaagcca | ttctgtctaa | gtgacagctg | agcagaacga | tgagtgagg | gggcaggga | 79260 |
| gaatgtggag | gtctggggga | agagcattcc | aaacagggat | ggcaggtgca | gaggccctga | 79320 |
| gggggcagca | tgcctggtag | gcctgctgga | tatttctgag | aatgaaaaag | tccattccca | 79380 |
| gttgtttgcg | tggagaacct | aaccagttta | taaacacaaa | gtacttgaga | ctgtggttaa | 79440 |
| tatacagcca | gctttccagc | agtgcacta | ctgtgtacat | caagggaaaa | ctgaacttcg | 79500 |
| ttttccttaa | aacttatcat | cagctgggtca | tcattttgac | aaattctgtc | aacaacagca | 79560 |
| gtgtcatctc | tggcatctgt | atgggtcacg | tctgaacaga | cacacgccct | cgagccctgc | 79620 |
| aggtaccagc | tgtataacaa | gaactccctt | ccacccctgt | tccctggaaac | aagaaagcca | 79680 |
| ttagaccgga | agatcccgat | ggctatctca | aatgtgctgg | atggagttgc | caggggccac | 79740 |
| tggcatccoc | tgtaaagcct | tgtctccacg | tttggttcc | ggcccttgaa | gactccattt | 79800 |
| ctgagtttgt | gtgtgtttta | ctttctagt | tgtgtcctca | tcttaatttt | tctctctctc | 79860 |
| ttctgccttt | aactgaaggt | tcgcttgggt | gtggagagac | agggcccccag | cagagcagct | 79920 |
| tcccgagaca | tctctcgatc | cagggcttcc | cagcagcccg | gcaagcgagg | gctgtgcctt | 79980 |
| tctgcttaag | ctcacaagca | tgccaggctc | actggcaagc | tgctgtctgg | ttgagggaat | 80040 |
| gctcctaag | ccctgcacag | ccctgtctct | cctggccctc | tggaaattcc | acccccgtgt | 80100 |
| ccacatttca | tgcaaaaatg | agctggttct | gtgagcatgg | cccgccctga | ctcgttagt | 80160 |
| gggcggtaag | tggtttccac | ttcaaccttg | cacctaatca | ccgggtctca | caccaggtat | 80220 |
| gaactcatag | agccgtgaag | tttccagtta | taaatccaca | gagcttcca | gcacctgcct | 80280 |
| tttcgcatca | cctccactcc | cagccacctg | ccaggcaaca | ggttaacagag | accaggtcac | 80340 |
| aggagggcag | tgtgggggca | ggactgcagt | ctccaaaagc | ccatgcacaa | aaccgacagc | 80400 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gccctggcag | gacaaggagg | ctgacattca | gatgtggagg | aacaaggcat | gacccattcc | 80460 |
| tggtcacggg | ggccacagct | gggactcagcc | ttgaggcttg | gccagactta | acaccgtgta | 80520 |
| taaacaggga | ccttttttagg | tagagtaatg | gaaaccaaac | tctaattgatc | ttagacagtg | 80580 |
| ctattagctc | cctggagctg | ccagaaacaaa | ttaccacaac | ttcaggtgtct | tgaacaagga | 80640 |
| gaaactgatt | ctcacagttc | tggagaccag | aagtctgaaa | tgcagggtgt | gccagggtcg | 80700 |
| tagtctctgg | agactctagg | ggaatctgtg | cctacotcct | ccagcttcca | gtgggtcctg | 80760 |
| acattccttg | gcttgttgct | gcatcacccc | aatccctatc | ctgtctctcc | cctgtctctt | 80820 |
| tgctcaaaat | gtctgtgttt | agtttccctg | tagacacctc | tgcatacctc | tcataagatg | 80880 |
| cagagggtgc | acatacaggt | gttgagagcc | cacttagata | atccaggata | agctcctctc | 80940 |
| aagatctgta | acttggctgg | gtgcagtgct | tcacacctgt | aatcccagca | ctttggggagg | 81000 |
| ccaaggcggg | aggttcactt | gaggtcagga | gttggagaac | agcctgggca | acatggggag | 81060 |
| accctgtctc | tactaaaaat | agccaggcgt | ggtggcacac | acctgtgttc | ccagctactc | 81120 |
| aggaggctga | ggtaggagga | ttgcttgagc | ctgggagttt | gagggtgcag | tgggctatga | 81180 |
| ctgcaccact | gaattccagc | ttgggtgaca | gagtgagact | gtctcaaaaa | aaaaaaaaac | 81240 |
| ataaaacata | acttaaatca | catctcttgc | cacagaaagt | aatactcttt | tgctacata | 81300 |
| taaggtaata | tttacaggat | ccaggggtta | ggatgtggac | atatctttgg | gaccactgac | 81360 |
| agccatgaag | caatctcata | attttcaaat | aggttctgtc | ctttttatct | ttccagctct | 81420 |
| ttggaaagca | tatgcctata | ttttcaatcc | acaattctat | ttttatttga | ggtcatttca | 81480 |
| tttctgtttt | tattttttta | ttgagacagg | gtctcactct | gtcacaggct | ggaatacact | 81540 |
| agccaatca | tggctcactg | cagccaactt | ctgggctgaa | gtgatcctcc | agcctcagcc | 81600 |
| tccctagtag | ctggaactac | agacacacat | caccatgcct | ggctgattca | ttttttaaht | 81660 |
| ttttctagag | acaggctcta | tgttgcccag | gctggtctca | aactcctggc | ccatgcgaaa | 81720 |
| cctccgcctt | cggctcccca | atgtgctggg | attataggag | taagccgcct | taccagcctt | 81780 |
| ccagttttat | tgtgttttgt | tttgttttgt | ttgagacaga | gtcatgctct | gtcaccaggc | 81840 |
| tggagtgcag | tggcagcagt | tgggtcactc | gcaacctctg | ccttgggggt | tcaagcgatt | 81900 |
| ctctgcctcc | agctcccgga | gtagctcgga | ttataggcat | gcaccaccac | gctggcttaa | 81960 |
| atttttgtat | tttttagtaga | gaccgggttt | caccatgttg | gtgaacacaa | agtatttgag | 82020 |
| agctgggtta | atatacagcc | agctttccag | aaatgcaact | actgtgttca | tcaagggaaa | 82080 |
| actgaacttc | gttttcccta | aaactttatca | ccagctgggtc | atcatatttga | caaatctctg | 82140 |
| caacacacagc | catgtcattc | ctggatctctg | tattggtcac | atctgaacga | cacacgcctt | 82200 |
| gcagtcagcc | ctgcaggta | tggctgtata | acacgcctgg | caagaaattcc | cttccacctt | 82260 |
| gtgtcccgga | aacaagaaag | tgatttgtat | ccactcacct | cagctctcca | aagtgtctggg | 82320 |
| attacaggcg | tgagccactg | cactcggctt | cctcccggtt | tttttttttt | tcaatgctta | 82380 |
| tattttactc | taattaaactg | agtcaaaaat | tgagaatagt | tgaatacact | ctcatgtaag | 82440 |
| gcgaatcatt | tagccgatac | ttaactctgc | atttgggcta | ccatgccgct | gtggtttagcg | 82500 |
| gggcaagggt | atgagccctg | tctcaacaca | cacacccgc | ctctccccag | cccacttaca | 82560 |
| cgcgtccatt | cccacgcagg | tgtggggggc | ttagaggatt | ccctcttctt | cgtaaaagtga | 82620 |
| gaatgggctg | gactcggctt | cactgcccaa | caactccttt | ttttctcttg | ggaaactgtc | 82680 |
| cttcccattc | catataacct | tcatgggctg | agatcactca | ccgagctaca | ggagaggggc | 82740 |
| catactatg | gccccatcag | ggttctgcct | gggacagtc | ataaaatgctg | gaagagagag | 82800 |
| gtgcctttcc | ctactgagcg | tgtctaaaagg | agacactgca | aactcggggc | gctgggtggca | 82860 |
| atcttacact | ctcagtgaaa | cctgcctcgg | ctgcagggga | aaccaattgca | gagccagca | 82920 |
| gggggcaacat | catttgaaac | gctggagaca | gctgtgcctg | aagccacat | ggctcaggga | 82980 |
| catgctcaag | accctctgat | ttgcttgta | cagtcaagag | agtccttggc | cgggacagat | 83040 |
| ggctcatgcc | tgtaatccca | gcactttggg | agggccgaggc | agggcgatca | ctcaggttca | 83100 |
| ggagttcaag | accagcctga | ccaacatggt | gaaaccccg | ttctactaaa | aatacaaaaa | 83160 |
| ttagctgggg | atggtggcat | gtccctgtaa | tcccagttgt | tagggaggct | gaggtggggg | 83220 |
| aatcgcttga | acctgggagg | tggaggttgc | agtgagccaa | gattgcacca | ctgcactcca | 83280 |
| ggctagctaa | caaagcgagg | ctctgactca | aaaaaagaga | gtcctgactg | aagctgagag | 83340 |
| gctggtggag | agctgtcagc | aagcagtggt | tgtgagctct | atgtgagctg | cgaagctcac | 83400 |
| acctgatttc | agagctgggtg | gctcttttct | caatcaggac | agctccaggc | ctgagttttg | 83460 |
| gtgtggtctg | tacctaaacg | ctgtgtctta | ggtcaggatc | tctagaagcc | aagccccaaga | 83520 |
| caggagtctt | caaatgcogt | gatttactga | gggaatgcct | tcggggggaac | ctgccagtga | 83580 |
| ggagagctgg | aggagcgggg | cagggggggg | agctgagcca | ggtggggggtc | cctgtctggag | 83640 |
| tctagctcag | gcttgacccc | acagggggctc | tggagcagga | acagcacact | gcatgtctcc | 83700 |
| ttgagcgat | cactggctgc | agttgcctct | gggagcagag | taaaagtggg | caggcatttc | 83760 |
| tgggagtaca | attccctaga | gaaggggaca | gctctgagct | gtgttaccag | ccaccatttc | 83820 |
| caggggctgg | gggatgcgct | gcactggccc | agagaggtct | ctgagcaagg | ccccacaaac | 83880 |

| | | | | | | |
|-------------|------------|-------------|-------------|-------------|-------------|-------|
| ctccacttca | acctcctgcc | aaacctcaag | cagattgaga | aggagccgct | agagacacag | 83940 |
| gaacctttc | ctcatcagaa | gtctccaagc | tttggaggaa | gaaggaggca | ggcaggaggc | 84000 |
| gggaggatc | tttcacccca | gggggttcagg | gctgcagtg | actatgatca | tgcacttgca | 84060 |
| ctactctctg | gtgacagagt | gagacctgt | ctcaaaaaca | aacagacaaa | aaaaatcaat | 84120 |
| gctcttagca | tctgctgggt | ccacattggg | cagtgctgt | gacgtgggtcc | agctgcaagg | 84180 |
| gaggttgagg | aatgcagttt | cattatctga | gtactagtgg | agaaagaaga | gacgaaagca | 84240 |
| atgagaatcc | ctgccatgag | gaagttagag | ctgattcatt | tcagcttcct | aaggtgaggt | 84300 |
| gatggggccc | agacaggttc | tgaggccact | ggaggtcaca | caggccagca | gggatacacc | 84360 |
| cagtggtaga | ccctggcttc | caggtcccca | ggacacagac | cacataaaag | caaggtcagt | 84420 |
| ctcagcacac | agatttaata | agcactgtat | gttatgccta | cagctgcagg | aaaccagggg | 84480 |
| gaagatggcc | acagggccaa | tgccaggcag | ggcacagggt | ggggcccttc | aagcagcata | 84540 |
| tggggaacag | gaaagactct | gccagcatgg | tcattggtgca | cagggtggcta | caatggtggg | 84600 |
| taatgcccg | ttgggaggat | gcagtggttag | acggggtctt | gtccgatgga | ttccttaggg | 84660 |
| ggttatatgc | ttgggagggt | gatacaattt | cctagaattc | tgggatctct | gatgtctctc | 84720 |
| gacctgtct | aggggatctg | ggtctccaga | gaggtctcag | ctacaaaagt | gacctctcc | 84780 |
| tgacccaccc | ttgtgggaga | ccccagtggt | ctcaaggtct | gtgaataaag | ggacttacaa | 84840 |
| ccaagatgcc | tctctgacag | gagacggggg | ctttacagtc | aggcccccagc | ctacatccct | 84900 |
| tgcctcagg | gcgggcacac | aacctacaaa | gggcagcgga | gtgaaagcac | gtgaaagcac | 84960 |
| cagggccggc | ttggggggta | agccagcccc | ttctcaagg | gccagtaggc | cacagaggcg | 85020 |
| gcccaggcag | gcagggtgga | gtagctggct | gtgctatccg | aggtcgctgt | cctaatacaga | 85080 |
| gcagggccgg | gagagccagg | acaggaagta | tggagagcag | ggagcgtctc | tccagggccc | 85140 |
| tgccttgtag | ggacaccttg | ggggtcggag | ccaagtgcag | acttgaggca | aaagcgtctc | 85200 |
| tctgcctct | gcctgcaccc | cccccatg | ccctcatgc | ctccaccagg | gcccaccagg | 85260 |
| gacggaagg | tgtgaaggtg | acagtggggg | aggtgccagc | tcacctctcc | ctcctgcac | 85320 |
| ctacctgcag | gaggtctg | tccgggtctc | cctgggggtt | cagctggctc | agcaggactg | 85380 |
| ggggcccaacc | gctggcaaa | gcctgaatgc | caccatctat | ggagggagg | gtccagggtg | 85440 |
| gacttagctg | ggaaccccta | gcacagggtc | cagcacagg | gacaaaagta | gagaccccc | 85500 |
| gctgcгааа | gtggtcctt | ggacagaagg | tgtgtggtga | ggaataacaa | aataataagc | 85560 |
| acagctgtga | cagagtga | caactgtaca | gtcttccatt | ctccccctta | aaatggtctc | 85620 |
| cagccagggc | gcttgggagg | gggcagtaag | cgccgcccc | actccccctc | cactcccccc | 85680 |
| cggggccctc | acccaagcag | cggttcctgg | taaagagccc | cgggaaggct | tcgcagctct | 85740 |
| cacgcgggtt | cccgtctgct | ccgcagtcgc | accagggcgc | cacgcgcgcg | ctcacgttgt | 85800 |
| ccacgtagtt | aggggtgacg | gcggtgctgc | cggggaccct | gagggcgaa | tcactgcgcc | 85860 |
| accgcggcgc | agatatcccc | tggagaaccc | ccgccctcgc | ccggatcccg | gcgcgcgcta | 85920 |
| ccacagaggc | ccgcgtaggc | gcgcaggcag | cgggcgccct | ggtccagcag | gcagcgtctg | 85980 |
| ggggcgctgg | gcgtgggggt | gcacgagacc | tgaaggccca | ggagggcagg | cctgcgcggc | 86040 |
| gggagggcgg | tgaagccgag | agaggcccg | tccccacct | gcgccaggcc | ccgcgcgcgc | 86100 |
| cgcgcgcac | ctgcagaccc | ggctgcgctc | gcgaagttt | aagggtctca | ggcaggagg | 86160 |
| cggcgcggg | ccgggcccc | aaaaggcgca | ggagggcacg | aaggtctggc | gccgcgcctc | 86220 |
| ggcgcacgcg | gggcccgcgc | acgggacagaa | gagcagtcgc | tgggttgagc | cgggcggccc | 86280 |
| gcgggcgaa | aagcggcgca | gggcccggcg | gcagcgggcg | cggggacagc | ccctctgcgc | 86340 |
| agcccgccc | aggcactgcg | ccacacatcg | ggagcgcaaa | cgtctggcacc | gcgcgtccgc | 86400 |
| cgtgcaggct | tggcgccgct | ccacacatcg | gttccctccg | accgagctcg | ccgacctggg | 86460 |
| gaaaggcgcg | gggaggctgc | aggtctcagt | gcgcgccgca | tgccacactt | accagcccca | 86520 |
| tccacgccc | tggaaagtg | gggacactga | aaaccgcaga | aagcatagag | tgtggcccaa | 86580 |
| aggcggggcc | gactggcact | gatcgggtct | ccgtcgtggg | gatgggcagc | ggcgggtctc | 86640 |
| ggagggtctc | gcaacacgtc | agggatggag | agggagcatg | gaggtctgagc | caggccttcc | 86700 |
| ccaaaagctc | aagccctctc | ttcccaagca | tcgccaccat | tccagacaag | gccacacaag | 86760 |
| tccaaagacg | ggtccagggg | ctcctcctac | actcaatggc | actactcgac | gggggaacaa | 86820 |
| tgaggctttc | tggggatctg | gcctaaggag | actttaaac | cagagatggc | atagaagaca | 86880 |
| gtggctgctc | ctagtctctt | aatagacggc | tgtgacgcca | tacctttcct | cctctacccc | 86940 |
| ctccacctcg | ccgttatctc | tccacatctg | agacctttg | ggtgggtcta | tgtattctgt | 87000 |
| tttctctctc | ctctctgacc | tctctctctc | cctcctgccc | cctgcatggg | tgtctccctc | 87060 |
| ctccagccct | tctgatgctc | cccttgagtc | gaggtcatcc | tgccccatct | gcactggctt | 87120 |
| cctccagccc | ctccaccttt | catcagccag | gggagagcgg | atttatgcaa | gcactctggc | 87180 |
| cactgaagaa | cccactaggt | cctggtctgc | aggggttgga | ttcagagaag | tccctgacga | 87240 |
| caaccgcccg | ggctttgtcc | tgagccacag | gacccaagg | ttcatcagtg | gttttagtga | 87300 |
| ccaggcccca | tgataggtag | cacctgtctg | gagacagaga | aggtgtctgg | cacccagtc | 87360 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| acacaccagg | ggttgggtag | acagaggggtg | agctctgccc | acatgccccac | aggctgtgtgc | 87420 |
| taaatagagc | caggataaact | ggaagttggg | acttggcagc | cccttcccag | gattttagtc | 87480 |
| acaaaaacca | ggctaatacga | atccccccca | gcctccacaa | gtctccagtc | ctggagcacc | 87540 |
| tcattcagac | tagaagcctt | tctggttgccc | agcctgagtg | gacgtgcaca | gcttggcccc | 87600 |
| tcattccatct | ccctgtgtgtg | ttgtgcacag | gggctacggg | gtggatatccc | ccccaggggc | 87660 |
| tgggagcagg | gtgcctctctc | ggcaatggct | gtgtctttca | ggctgtggat | tggaccgcgt | 87720 |
| ggtagagagt | attcctgtctg | tgaactccag | acctgccttc | gcaggaaagg | tgaacatagc | 87780 |
| tcgaggggctt | agaatgggtcc | ctgagtgagg | gggatagatg | gcagcgggggt | gagtcctccc | 87840 |
| ttcctccccc | ccccaaaggg | cacactttgc | tttctctctca | cagccttgcc | aggcacatcc | 87900 |
| tcaggtctca | gctgcacaca | ccagcacctg | atggtgccta | agggtttgaa | cccactctgc | 87960 |
| cctgttgctct | acctgtcacc | cttgcccaca | acaccccttac | ccagtctcca | ccctccacagc | 88020 |
| agctcatctc | agctgcaaga | tctctccac | cttaaacctt | taccagattc | cactttccac | 88080 |
| actacactcg | gtaaagtgc | ccctccctcc | gcgacatcca | tgcagggaag | gagcagtcga | 88140 |
| agaagggtgga | acaaacagatg | aacaacaccg | aagattcacc | aaggaatctg | tgtgatgggt | 88200 |
| aggagggtcg | gggtctgagg | gaggttgaga | gagttgggag | atggcgacaga | ataggcaaaa | 88260 |
| gttctctggg | cacgggtgccc | aaggctggac | tcgggagggg | gatggcggtg | gtggcttgac | 88320 |
| ggggctggga | ggtagagtg | tcagcacata | gtggggtcaa | ttggccaggg | tccccaagcc | 88380 |
| gtgcgcacag | gccctctcact | ggcaacccct | cctgtgggtc | agcttttctc | ggaatatctt | 88440 |
| gccctcccac | cttgtcactg | aggagcagg | acagagaggc | cagggttcgg | ggtccagggg | 88500 |
| aagagagggg | cgctcaccca | gtaacagcag | cagcagcagc | gcaggcccca | ggcagcggac | 88560 |
| catgctggac | cttcaacaga | gaaggatggc | tggcagagtc | ctagcttgat | aggcgcccc | 88620 |
| ttcctagtgc | ctctggagca | ggagcacagt | gaaaacccag | ctgtctgcgc | tgaataccct | 88680 |
| ggggaggagcc | tttgtctcag | caactcccgc | cctcattaa | gtctccacc | caggctgcac | 88740 |
| cgtcaggtaa | acctgggagc | cacgggattc | ggcgccctgaa | tgtgcctaga | cggagcccaa | 88800 |
| cacactagcc | ctgcaccgca | caccaggcaa | cacacagaca | accaccagg | gagacctggt | 88860 |
| ctctcccaatg | ctaaactagg | aacgcatatt | ttcttttctc | ttcttttctc | ttttatata | 88920 |
| attctatttt | aaacttaaaa | aaattttttt | gaatagagtt | ggggtctcgc | tctattgcgc | 88980 |
| aggctggtct | tgaactcctg | ggctcaaaac | atcctcccac | ctcggtctct | caaagtgtca | 89040 |
| ggatgacagg | tgtgagcac | catgctggc | ctattttaaa | ctctctattg | ttgactaatt | 89100 |
| ttagacttgc | agaaaagtgt | gaaaatgtaa | gcattttcct | tacacccctc | accagcttc | 89160 |
| ccctgatggt | ccatcttact | taatcatagt | ccaatcatta | cagctaggaa | actaaccttg | 89220 |
| gtacaatcct | gttcaacaaa | ctgtgactgt | ttttgcattt | ctcatitttc | caactatgtc | 89280 |
| ctttttctgt | tccaagatcc | attctaggat | cccacatcac | agttgtctcc | ttattctctc | 89340 |
| caatctggga | gttccctcat | ctttccctat | ctttacocct | gacacttctg | agaatcctcg | 89400 |
| gccagttatt | ttgcagaatg | tttctcttga | gttgtctgat | gttttattct | gatcagaatg | 89460 |
| agacacagca | ttgttttgac | taacccaaaa | gttattctat | aagtaaatat | tgagggtaaa | 89520 |
| aatctcatcc | aacctgggca | acagagtaag | gccctgtctc | aaataaagtc | tcaacactaa | 89580 |
| gatttaaaaa | gtgacagaaa | aagccccac | tatgatttgt | cttgactttt | ttttaaaaa | 89640 |
| caaacaaaa | aacaaaaaaa | caaacaaaa | cgatgtcttg | ctctctctct | ccctcagggt | 89700 |
| ggagtgagct | ggcacgatct | tggctcactg | caacctccgc | ctcccagggt | caagcgattc | 89760 |
| ttctgcctca | gcctcccagg | tagctgggat | tacagacgcc | cgcaccatgt | cccggtcaat | 89820 |
| ttttgtattt | tttagtagaga | caaggtttca | ccatgttggc | caggcttggtc | tcgaacttct | 89880 |
| aacctcaagt | gaccgcgcca | cctctgcctc | ccaaagtgtc | ggaattacag | actgagcaaa | 89940 |
| ctgcgcttgg | cctcatttagc | atcttaaatc | tccacacagg | gtgtgttctc | ttactgttat | 90000 |
| aaggagcaaa | ggatcagttt | gaggacaggt | aaaataaaaa | tcgccttgct | gctagagggg | 90060 |
| agaagtccct | gctgaagata | gcttttgcttg | aatgagctca | attgcaatgc | cagtgctgag | 90120 |
| gctgtgtgac | tgtacgggtca | ccacaggttg | tgtctgcgcg | ctagaacatg | gtcactttct | 90180 |
| tgcataccta | tctgtctcca | gtacatctgt | ctgtggtttg | tgggtgtcca | tttcttaatt | 90240 |
| tttttaatga | atcagaagac | tgtgatgtgc | tttccgctgt | gctaaccatg | gccgctgaag | 90300 |
| caaaatgtaa | accaagatgc | ccctgcagtg | gttgtgcttc | actctacgac | atctgttacc | 90360 |
| ggaaaggggt | ccagattcag | acccaggag | agggttcttg | gactctgctg | agcaagaagt | 90420 |
| ttgagacgag | tccataaagt | gaaagcacat | ttattaagaa | agtaaaagaa | taaaagaaatg | 90480 |
| gctactccat | agagagcgca | gccctgaggg | ctgctggttg | cccattttta | tggtgttttc | 90540 |
| tgtgatctca | gctaaacagg | gggtgattgt | tcatgtctcc | ctctttttaga | ccatatatgg | 90600 |
| taacttctctg | atactgcgat | ggcatctgta | acctgtcatg | gtgctggttg | gagtgtagca | 90660 |
| gtggggaccg | accagaggtc | acctcatatc | ccatcttggt | ttgggtgggt | tttagccagc | 90720 |
| ttctatatgt | caagctgatt | tttttggttg | gttttggttt | tgagacggag | tctcgtccca | 90780 |
| ggctggagtg | cagtgacatg | atctcagctc | actgcaacct | ccacctcctg | ggttcaaggc | 90840 |

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|-------|
| attctcctgc | ctcagcctcc | caagtagctg | ggattacagg | cacacaccac | catgcctggc | 90900 |
| tagtttttgt | atttttagta | gagatgaggt | ttcaccttgt | tggccagggt | ggtctcgaa | 90960 |
| ttctgacctc | aggtgatccg | cccacctcag | cctcccaaa | tgtctgggatt | acaggcggtga | 91020 |
| gctaccgcgc | ctggcctaact | gcaaaactatt | ttatcagcaa | ggtcctttatg | acctgtatct | 91080 |
| cctatctcat | cctgtgacgc | agaatgctgt | aactgtcttg | aaacgcagcc | cagtaggtct | 91140 |
| cagccttatt | ttgtcagcgc | cctatttcagg | atggagttgc | tctgggtcacc | aggcctctga | 91200 |
| cacatctctc | tgtgttttgg | cggtgtggag | gaaagagggg | tgagggaagg | aaactcaaaa | 91260 |
| ccaagctctg | accacacagg | gcaggtacac | tctcccaact | gtctgtgggt | gccacaagtc | 91320 |
| aaggggaggg | cagagagaga | agaagggtgt | acagatggcc | gcaggccaca | gaatgtcaga | 91380 |
| ggaagcccag | ttcctcccgg | ggcagcccaa | gtagctggta | gttgggtggc | caaacagagg | 91440 |
| gcgtcacagc | tgagctgggc | tcgctcgcta | ccccagctc | agcgtccact | ctgcacctca | 91500 |
| tgacctctg | ctcagctca | gggtccatgc | ctaccctcct | gcttccag | ctctctgcg | 91560 |
| gtcctctg | ttttctgtg | ttggcccat | gccagctcct | ttctgtgtg | ctctctctct | 91620 |
| ccagttctg | agcacagcca | gggtgatcctg | ggctccagac | aggcctctcc | ccagctctgg | 91680 |
| ggctccctc | cttagagccc | tctttccttc | ccacgtggcc | tccccgggt | tccgactga | 91740 |
| atggagaagg | ggtggagggg | gtgctgggca | gtcctgggga | ttagccaaga | gggcagaggt | 91800 |
| ggctccccc | gggtcccttg | tagctggagt | ccgcgggggt | ctagacaccc | ctctctgaag | 91860 |
| ggtaagagcg | ggggaggtat | atlaacgtgt | atttttagag | tctctctttt | ttttttttt | 91920 |
| ttttttgaga | cgaggtcttg | ctctgtgtgc | ccaggctgga | gtgcagtggt | acgatctcag | 91980 |
| ctcagtcgaa | cctccacctc | ccatgttcaa | acaattctcc | tgcgcagcc | tccgagtag | 92040 |
| ctgggattat | aggcacgtgc | caccacaccg | ggctaatttt | tgtattttta | gtagagatgg | 92100 |
| ggtttcacca | cattggccag | gctggtctcg | aacttctgac | ctcaaattgat | cctcccgct | 92160 |
| tggcctccca | aagtgctggg | atgacaggcg | tgagccacca | gctctggct | tagattctct | 92220 |
| atttttaacat | ggtttaacat | ctcggggtgg | gggcttgttg | gtcggagaga | aactgcttga | 92280 |
| ttctgggaga | tcagaaacaa | ctcatgctct | tcatatgcaa | accgaccagt | cttgagttcca | 92340 |
| tacaccaacc | accccttcca | aggaaacttc | acatacga | ccagattttc | cctgcctcta | 92400 |
| aaccagctca | gggccaggca | cctaccaga | caattagagc | ccaacccac | gccccaaacc | 92460 |
| ccacccaga | gattcaaaa | gccaatccta | ccctcttccc | tgggcttgc | ttctctcccc | 92520 |
| agtggaacgc | gcactgtggg | ctgtggctgt | tgcttccac | ctgctctgta | ttctgtccct | 92580 |
| ggacaaaacc | tagtgctctc | ccactgtggc | cctgcatggt | gggaaactgt | gagtaataac | 92640 |
| ttattttcaac | agcattggcc | tctgtgtcat | cagtcacctt | cataaattaa | aatttttgcaa | 92700 |
| ctcaactga | ggcaggaggg | ggcttttagag | agcactctca | cttgcccct | tgggcttgtg | 92760 |
| gagtgaggcc | cgaggtgaag | ttgctctcct | gcactcagtt | ttgggatggt | tttggtatgc | 92820 |
| ggcaacagct | ggctgatata | ggccactcag | cagctcttgc | atgaagcaat | gggagaaatg | 92880 |
| ggaacgccc | gggagggcag | agtgacagag | caaagagggt | gttcaaaacta | ggcaacaccg | 92940 |
| tcctgtgtccc | agacacatgc | ctggggctct | ggctgccatc | tagggctcggc | ctgccagcca | 93000 |
| ccaccccgtc | agccaccacc | ccaccaacca | ccaccccgc | aaccaccacc | ccgtgtggtc | 93060 |
| ctcagggcac | cccatgggtt | gaatttacat | acagaaaaga | ctaaccaagg | tccaagtgt | 93120 |
| atatgtcttt | ttaaaattta | tttttagagac | agggtcttgc | tctgtcacc | aggctggagt | 93180 |
| gcagtggtgc | caccatagct | cactgcagtg | ttgaactcag | gctcaaagtga | tcctcctgct | 93240 |
| tcagcctcct | gagcagctag | gcctacagg | acacaccacc | acgcccagct | aatttttaaaa | 93300 |
| tttttctgta | gatcggtgtg | cttgctgtgt | tgcccaggct | gggtgtggac | tcttgccctc | 93360 |
| tgtgtgatcct | cctaccaccag | cttcccaag | tgctgggatt | ataggcatga | gcaccacgc | 93420 |
| ctatagccaa | catgtctctc | ttttgacttc | tacttttggt | tcttttctga | aatgggtccc | 93480 |
| tctgtccccc | cgacacacac | agaatggggg | agaggctgtc | agattctbag | ctccagaacc | 93540 |
| tcagggtgat | cactgggatt | gggggtgggg | gctcaggaa | cacctagggg | agaagacagg | 93600 |
| gtggggaagaa | acaggaagga | agggtcccaa | aattatgttt | gtttgcagag | gccagccagg | 93660 |
| ctccagggga | gtgtggactc | agtcgaacca | tagggcccca | ggaccactag | cttctggcca | 93720 |
| gcagtcabgc | cctccacaga | gctgggtccg | tggaaattgc | atgtaggaga | cacaccagac | 93780 |
| tccagggaca | gagccctttt | gggatggcca | gcactaccca | gcttccactg | gtgaggagg | 93840 |
| tcaggggctg | tgtgactctt | ctgtctggga | ctgatgggtt | attgagctgg | agagtgtgct | 93900 |
| cagcagtggt | ctccagccct | caggaaactc | tagtgtggct | ctgggttccct | ggagtgggtg | 93960 |
| ggtcgaagct | ccactcgggg | aagaaacttc | caagctgcct | gcaggtgtgt | gaggtccggt | 94020 |
| gattcaactg | ctctgcccct | gcagttcaag | ttcctggagt | ggctctcagt | ggccacctgt | 94080 |
| cttttaactct | tgtcatttta | ggagctacct | ctcaccagag | gcaggtatct | ggcatctgga | 94140 |
| cttgatctgc | tgagaatgag | gaggatagt | tgtcccctaa | ggaactgggc | ccagggtgc | 94200 |
| aagctgtgtg | gcagagagcc | catcctcact | cagtgaggac | cagtgatcca | ggaaaagcca | 94260 |
| cagctctctc | ctccccagcc | caggggtctc | cagcatcctg | gtctccatga | taaccaagag | 94320 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gtcataaact | catttccata | ataacctgag | cccagaaacc | tgattagggg | gcagcaaaact | 94380 |
| gaggggtggg | agaggtggga | gggtggggcga | tgagaggggg | aggctttgaa | tccaggtccc | 94440 |
| tgctactcctt | gggggttcagg | cgagactgct | ggcagagggt | tctcagggggt | gctgctgggg | 94500 |
| tcatgagagt | tctcagggtc | tgggagaaat | ggtggagggt | aaatgttgtg | aatatgggtca | 94560 |
| gcaggagacc | ctggggctgtg | ggaggggcat | aggggactca | agggtgactgg | gtgctgcccc | 94620 |
| tctggaagga | gcaggaggcg | atgagccctt | cccttctccc | ttccctctcc | acctccccct | 94680 |
| gggtgcctcac | tcacccagggt | gccaggggct | tccagtggtg | ctggggccca | actccaatggg | 94740 |
| gtgaacgcgc | cccaggggggt | ggtccctgtg | tgggccatct | tggggctga | gcaacgtgat | 94800 |
| aagagtcacg | gaggttgga | cagtgatcct | gagtggggta | ttgcccctccc | gcagcatggt | 94860 |
| gtccagccca | gggagttctg | cgtttactga | gtttcttggg | gcacccatct | gctccaagtc | 94920 |
| acctctctag | ctcccttctc | gctccctctt | caggggagcc | ttgggatcca | gctcccaagt | 94980 |
| gagccctcatg | ccctcggtcg | gcacctctc | tctctagtcc | taacatttcc | tccaggctct | 95040 |
| gacacacccc | agcagcctgg | cactctccag | atgctggcat | cgctcagctt | ccaaagaaac | 95100 |
| ttggatgtcc | gccccctcgg | cagctatgtc | tgtctctcct | gccccctggg | gccttgctgc | 95160 |
| ccttgatgat | tccaagccat | ctttgactgt | ccccatcca | tcccccaagg | ccttgctcatt | 95220 |
| tcctgtgatg | ttccttcaaa | acattctccc | ctgccctgag | actcccgcct | ggggatgaga | 95280 |
| agcagcgccg | acccctctga | gcgccccctc | cgctgctgca | cgccagggtc | ttggccacctt | 95340 |
| gctcctctgc | ccacagaccc | tcaatccaaa | ctcgctttgt | cagggtctctc | ttagctgccca | 95400 |
| cccggggccc | aggtgggtgcc | ctgcccctgt | ctgttatgoc | ctctgcccc | atctctggcc | 95460 |
| caaatcatgc | catctccctt | ggcttgcccg | gagcactccc | aagaccaggc | tatgtcagac | 95520 |
| atggccacag | agtgctgccc | ctgcttaggg | ccctgggtgca | gggtgagtc | taggacagcc | 95580 |
| atgcttagta | ttatgtgact | ccccactccg | ccaccaccca | ggtcacagag | aaactgggtta | 95640 |
| aggcaggggc | ctggcacagg | ggcagccagc | accgcagctg | accagtggtta | ttgagtgaaa | 95700 |
| agatgtgctg | ggcccagcat | ttgggaactt | caaggggggtg | acagaggtga | tgtgtgcaga | 95760 |
| gggaagtggca | aagggccgaa | aactgggtgag | acagagggtg | gacaggcctc | cggggggcagc | 95820 |
| atggttacagg | gactgcaatc | tgagccaggg | aaaaacaggg | cgaagtcaag | ggtgaggcag | 95880 |
| ccagctgggtg | ggagaagcag | gagagtggac | aagaggagct | gtactgggag | gtagagggcc | 95940 |
| atgacctgag | gtgctgggtg | tggggcaggg | atccaccccc | tcgcttgact | gggagggcac | 96000 |
| tggaacctcc | tgttcaaaag | tactcttttc | catggcctct | ggggctgctc | ttcgacctg | 96060 |
| ggggcagggc | tgaggggctg | ccccagctcc | cacagcccca | gcagagctct | gaggagggga | 96120 |
| accgcaggag | taggctcagg | aagcaggcgc | tcggagccta | cccactgcac | gcaggggtccc | 96180 |
| ttctgcagcc | ccagctgcac | cgctgcagat | gggctcctgg | gagtgctgag | caacaccagg | 96240 |
| ccagggccgcg | ccctggggagc | agaggcagca | ggacgctgag | gagcatggcc | agcagggaag | 96300 |
| tgtcatggtc | tgcggggatt | gggggaaggg | gcgctgagtc | ctgacagagt | gcaccacccc | 96360 |
| agctcctgcc | cacatgcccc | cccactggcat | actttcagcc | tgcacagggc | cactgtccat | 96420 |
| gctgccacca | aagcctggct | tgtcacagaa | gggtggagcc | cagcctggag | cacagtggca | 96480 |
| gttatggttg | ctatttgaaa | cctgcagaga | agagaagagg | agggtcacgt | aggattagga | 96540 |
| accccaagggt | cacccccact | cctcgggctc | tcaccccgtg | gctgtggcag | gcagtcaggc | 96600 |
| agcgctgaag | ctcctgggaag | gcattcttcc | tgcagcgctt | gctctggcac | acctacgggc | 96660 |
| agtgccaccag | gcagtgaggg | ggacactggc | ctgcgggat | caaacggcga | ggaggggtcg | 96720 |
| ggtgggcaga | gctcaccatt | ctaggctccac | actgggtgcc | tggctctacc | aggcccaggc | 96780 |
| caagcagggtc | cagctgggca | ctggggagtg | ccaaggctcc | cgacaagtc | acttcttggc | 96840 |
| catctagggtg | aacggtagag | tccactggca | ccatgtgcgg | tcgcagcagg | cttgcctttc | 96900 |
| cacctgggca | ctcgagcttc | ccacacaggg | catccctggg | gaggaagtag | aggggggtca | 96960 |
| acagctgcag | tacccccctt | ccccaaacc | actccatagc | ttctgtctcc | tccactcagc | 97020 |
| tccactccct | acctccctgc | acagggcagg | aagtggccct | cgctgtcctg | gccgcagttt | 97080 |
| cctgagcat | ctcccgcaga | gttcacaccc | tggaaacagg | cctcgggagc | tgggtgggag | 97140 |
| cctgaggaag | catggggcag | gctcggggca | gctcggagag | gggctgcgt | cagcgcgctt | 97200 |
| cagtttcccc | tgcccatctt | ccccacagta | gaaaactggc | cccaccagag | gatggggggc | 97260 |
| aggggtcgaag | gggtgctcgtg | tcctctcacc | agggcccccag | agctctggcc | actgctgctc | 97320 |
| cagctgtggga | catcgcccat | cccagcagta | gccactgccc | ctggcacagg | gtgagccgtc | 97380 |
| cagttaggtaa | acgtctgggg | gacagtgggg | ggaggtgccc | gtgcataaac | cagggagggtc | 97440 |
| acagtcaccc | atggcctggc | ggcacagcgc | tccagccggc | ttcagctgcg | caggtgacgg | 97500 |
| gtggttgggg | aggcagagag | aggccacgtg | cagtgagagg | tccatgccga | gagcgcggtt | 97560 |
| cggagctggg | ggagccaggc | ctacccaagc | ccagcaccca | acgggggaac | ctgagggcac | 97620 |
| caatcaacta | aggcccaacg | cccggtccc | aagctccccg | aaacctcac | ctgaaacctt | 97680 |
| ccatgccctc | accaggcagc | gcacgcagca | gtccccgtgg | gcgcactggg | ccccggggcg | 97740 |
| cagcgagcag | ttgtgagcaa | agcagcagag | gtcgcggcac | tcttggggacc | agaaaggcaa | 97800 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| gaaggggccc | ggtgagggcg | cagcgcccca | gacctgagcg | gagaggggcaa | gtggggggccg | 97860 |
| ggcgagccga | cttaacctgg | ccaggggccc | agtcacactc | ctcgcccgct | tcacacgaagc | 97920 |
| cgctcccgca | gagcgccggc | ggcaccggga | gtccgggggtc | cggggcattg | gagagggcaag | 97980 |
| cgccgcccc | cttgcggaag | aaggcgcgca | gctggcgggc | gctgcaggcg | ctgaacacgc | 98040 |
| gcggaacg | gtgcctaccg | gcacggggag | ggcatatgggc | atggaggggac | agtcccccaa | 98100 |
| cccccgcgct | tctctgatcc | ccaccctctgg | gcttgggtac | agccgcgaga | cgcgcagagc | 98160 |
| ccagagaggg | gaagttaacc | gcgcgaagtc | acacaacaag | cggggacagg | gacgatggcg | 98220 |
| ccccaatagt | gagcagcccg | ggacccaagg | tggaaatcgcg | accgcagcgt | gctcctcccg | 98280 |
| gtgtaggagt | aacctcgcca | ggttactcgg | aaaataatct | tcataccgtt | gagaatccac | 98340 |
| tttgccctgag | cttcttccct | ttaagcctca | taaaccaccc | tgaagcggac | actatgatca | 98400 |
| tcttccccat | ttacagaa | aggaactga | gggacgacca | aagaaacgca | cgggaggaag | 98460 |
| ttccacggac | tagccgcccc | gccgcagccc | cgacccccca | cccgctgacc | cggtggccgc | 98520 |
| agccatgaac | cagcctccgg | actcggccgc | agcctccacg | cagcagccgt | cggggctcgtg | 98580 |
| gctgagccgc | aggctgtggc | cgatctcatg | ggccatgggtg | gctgcggcgc | cgatggggag | 98640 |
| ctccgagtgg | tctcttccct | ccgtggggag | cgggtcactg | cggccgtaga | gcctcctgtc | 98700 |
| tctccctcgc | ccccgcccc | ggggctcacc | gtgctcagc | ctcccgagct | ctcggcgccg | 98760 |
| cacatgcctc | gcagcggccg | caggcccaact | gtggcgccct | ggaagggcgc | gcctctgggg | 98820 |
| gcgagcgcg | gcgtgaccg | gcggggccgg | gaggtgagcg | cgccccaccc | gggacccgcg | 98880 |
| tcggggctcag | aggcaaccac | gtgagcagct | gcgcggagtc | gtggggccgc | tcgcgccaca | 98940 |
| gccccggcg | ccactgcagg | aaggcccaga | gcgtggcgtt | ggcgtcctgc | gtgacgcggc | 99000 |
| tgcggtcccg | ctcgttccac | acctccaggc | cggtcagcgc | ctcatgaatg | tcagagtcc | 99060 |
| tgagaagctg | agggcgaggc | ggggctgaag | ccgggacagg | gcgccccatc | gcgcgcgttg | 99120 |
| tccttcgtgg | ggcgcccttc | ctcttcccca | aaacccacca | gcacctgcct | gtcctgcgcg | 99180 |
| cgccaccccc | atcacccgtc | tctccccgcc | gcccccaacc | tggttccacgt | agttggcgac | 99240 |
| ttccaggaga | cgtctgttgg | tgtggttcaa | gtttcgggtg | cagtaacaag | actgggaagg | 99300 |
| cagaaatccc | gggtgcttga | ggggctgagc | tggccccatc | cctgaccccg | ccaacccctg | 99360 |
| gggtctctcc | tcaccagggt | gtggctctgc | acaatgtaca | gttccaggta | cttccgggtc | 99420 |
| ctcgcgctt | ctcgccctgc | ctcgcggagt | gcaaaatggg | accttagtgc | gaagctcgtc | 99480 |
| ggcttgagcc | atgcccccac | accgcagctc | ccagaaggaa | gtttaaactg | ttttctggaa | 99540 |
| cttctttctt | cagacttcaa | taaaaatact | gggactcgag | gcctgtgaat | tcctgtctct | 99600 |
| cttgatttgg | agggctatag | atacagcatt | cccactccca | tcctgatcgt | gcccctgacc | 99660 |
| ctgctctggg | gcacaccagg | aaggtctggt | atgcccgctt | tgttccagg | atccctgtgg | 99720 |
| ccacaggttc | ctttccagg | gagcagctgc | tcocatccga | agatctcgtg | gggttgagaag | 99780 |
| tccttgtagc | ccgggggtgg | ccagggacgc | agataaatagc | tggcatctct | gctgaggggtg | 99840 |
| atcaggccac | taggggtgag | aggggttagga | gcgggtgtga | gggagctctt | tccccatccc | 99900 |
| aggccagcc | tcctctccca | gagctcacct | catccagag | caggtgcaga | ggactaccca | 99960 |
| ggagtgggg | aagcccttta | ctcgcccttg | gtagtggcaa | tgatcctagg | gaggaagggg | 100020 |
| ccagccccaa | atctcagcca | gggctggagc | aagaggggca | agagggaggg | tgtggtaggg | 100080 |
| gctggctcca | accgcccctt | aggaatgcaa | ggaggagtgc | gggttaggaat | gggtggggggg | 100140 |
| tacotctggc | ggtgcattccc | agagcccatg | gaagcatctc | accgtgtggt | tgggggcccag | 100200 |
| caccactggc | tgcccatctg | ggccgtagtgc | ggtttctatg | tatcctgggg | ccagcagcct | 100260 |
| gctgagagg | ggtgttacag | ggaacactga | atccagcttc | ctcctgcctc | ctccaggatg | 100320 |
| tctccagcc | tctctcccta | aatgtctaag | gagcagcttt | atgagtgcga | cactcacagt | 100380 |
| gtgtcttagg | gaagggacag | gagcaatggt | gacttgctca | gatcagaaac | tcttgggggt | 100440 |
| agagggaagga | gccttggtga | tggcttagtt | gtgggagatg | tgaatatggg | aagcaccagg | 100500 |
| gaggacggc | gggaggagtgc | ggaatagggg | aagagtttgt | ggttcccagg | ggacctgcag | 100560 |
| caggcagcag | gatccacagg | atcgggaggg | gaggagtgcg | gagacactgc | cgaagaatgc | 100620 |
| gacttggagt | tggggaatg | cgggtgacct | ccccagttcc | cctgcctgct | tcctctcttt | 100680 |
| gttgggcatc | tggctgaccc | tcttgcoccc | acctgcccct | gatccttgaa | atattttctt | 100740 |
| cagacttcta | gacccacat | acctcccacc | tgtccttcag | tgattgatgc | tcacccctgt | 100800 |
| cctccagaga | aaacagaatc | gccactggcc | cacgctgctt | ccacctctcc | tgctctctcc | 100860 |
| acacccactc | cggtaatgat | tccacttcca | ggctccatct | caacaggatg | tttcccacac | 100920 |
| ggtggatca | atcataagtc | aatctgtcct | ctttaaagaa | aatccttaac | cccaactcac | 100980 |
| cttggccctca | ttaccttcag | accaaccgct | aatgatgggt | gcttcccccc | tcccaggcat | 101040 |
| tcacaccagc | gcgccagctc | tgccccctac | ccctgccccca | cacacaccgc | ccaccttagg | 101100 |
| agbtgaggt | tgtgaccac | ccatttgaa | ggtagggaag | tcgggaaat | atgggtggcg | 101160 |
| acagtggaa | tagagtttgt | tccctgtcca | tccgactcca | caggggagaa | taaaatacgt | 101220 |
| gtcaagtgt | cagaacagcg | cctgcggcta | agcactcagt | aggtgatata | tactgataac | 101280 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ataatctggg | tgggtttaag | agcctgcgct | ccagcccgga | caccaccccc | accagaccca | 101340 |
| aggcaccttc | ctgagcacag | gtctgcctgt | cccttcccca | gtctcaaaatc | tttggctctg | 101400 |
| gatggtcag | gacacttgagc | acccaaatgg | ccttctgtga | tctggccctc | ctgggactct | 101460 |
| ccaaattcat | tcccccttgc | tccccctct | gtagatggag | acottccagc | caggtcagac | 101520 |
| aaactgctgc | ccccaaagtgt | agccactgca | tctttttctt | ttttcttttc | tttctttttt | 101580 |
| tttttttttt | tctttttttt | ctcttttttt | gagacggagt | ctcactcttt | gcccaggccg | 101640 |
| gagtgggtgca | gtgggtgtgat | cttggctcac | cacaacctct | acctccgggg | ttcaaggcat | 101700 |
| tctcccgctct | cagcccccta | agtagctggg | attacaggcg | ccgccaccac | gctctggctaa | 101760 |
| ttttttgtat | tttttagtaga | gacgggggtt | caccatgttg | gcaaggcttg | tctcaaaactc | 101820 |
| ctgacctcag | gtgagccgcc | cgcctcggcc | tcccaaagcc | accgcactct | ggctccctgcc | 101880 |
| attcccttag | cctgggggtgc | cggctcatct | tttccctcta | ggatttcttt | agactcagca | 101940 |
| tatcttgcaa | atgtccacta | ggtggtgctc | actcatcgcc | agcagggagc | taacaagccg | 102000 |
| ctctgggggt | tgggagggcg | gaggtgcccc | acagcggggc | tgacagcttc | agcggctctc | 102060 |
| ttcagcctcc | agggagccaa | ccacaggcct | cgtgtgactct | ccctgtcttc | tgcacctctc | 102120 |
| ctgggggtct | ctgcccatcc | agccaccgcg | acagatctgt | gtcagtcctt | gcccccaac | 102180 |
| actgaccccc | tctctccagc | cctaccccag | cctggcactc | actggttctt | ctccagctca | 102240 |
| agcaggagct | ctcggccttc | agcctccagg | gccaccagcc | ccatgtctgg | ctctggagac | 102300 |
| tggggcaagaa | agagtgtgga | gtggcctccag | gctctctgcc | gtccaggagg | tggcacagcc | 102360 |
| taggtggcct | gtggagcccg | ctggggagga | agtctctggg | gagaacgtgg | gctggggagt | 102420 |
| cagcaggacc | ccccacatac | tatggagggc | gtggaggagg | tgagaacata | caaagatgtt | 102480 |
| cccaaaactca | ggatgtttgc | actcctgaca | acagccactt | ggaaggcgct | tggcacagcc | 102540 |
| tgccaggcac | accagcatcc | tccccagaga | ccagaggctc | cagaaaagggt | ccccctcccc | 102600 |
| ggccccgcc | tcttctttca | tggccagaag | gggcatcaaa | agcaggggaa | gacagagggg | 102660 |
| tgtcaggagc | attatggggg | catcgggtag | ccatggctcag | ggcctctcca | gagcctctgc | 102720 |
| tacctgagggc | ttggttccaa | atgagctgct | gctcatttcc | tatagaattc | aaatttgaat | 102780 |
| cctccacttc | caattttggc | ctcttccccc | gtttctctgg | gctctccagc | gctccagca | 102840 |
| gcccccgctc | ctccggtctc | gacacctgct | tcaactggacc | cacgaagtaa | acatggacgc | 102900 |
| catctccagc | aagagagcac | actggtctct | agctagggtg | caggagggtc | gcttggacgg | 102960 |
| ccagccctct | ctctctcccc | cctcctcttg | cgtctcccca | ccctgtggga | acaccccact | 103020 |
| tcccccttgt | ccactcagcc | tggctggggg | cccagagttg | gagccggccc | aggagcttcc | 103080 |
| tgggaggctg | ctgcgccttc | ggaatgttta | acccccgact | ccttttctcc | aaaaatgcac | 103140 |
| tggcctgggg | ccctgtccaa | gggtctcaga | gtctttggag | ggaagttctc | cttcgcaagt | 103200 |
| ggggagcaga | tggctccttg | ctccccggcc | acaggcccca | caaggcctcc | agcatgagct | 103260 |
| catgagcctg | gaatggcact | tgcctttatt | gggaaaaggtc | tgcaccggga | aaaaggccat | 103320 |
| actcgaggtc | cctgttctcc | tgcagccctc | gctatcttta | ctcttgcctc | cctggtaacc | 103380 |
| tgcctcttga | tatatacccc | tcatcttgaa | atgtgagttg | ttcctgcctt | tgggagggga | 103440 |
| tacctagcct | ctactcttcc | tctgttacca | tcttggcagg | cttctggggg | gcaggggccc | 103500 |
| accgggtggg | gaagcagagc | ccctttgggg | ctctcctctt | ggtcacagcc | caggccagac | 103560 |
| agacaggggg | gcccagaggg | agagtgaacc | cagtggtgtg | ccagccttcc | cctcctgggg | 103620 |
| atggggaggg | caatctcaaa | gctcaggcca | gtgccgtgct | tgaccagtgg | aatggggggc | 103680 |
| ttatgggcct | aggggatccc | agtgaaggcc | ctgggttggg | agctgtcggt | tctctggggg | 103740 |
| cctctcagcc | ttcatggcaa | tgtctccctc | ccttccctct | tgggtgattt | ggcacagtag | 103800 |
| gctgaaaaat | ccaaacaaag | agggctctct | aggaggggca | ggctgttagc | caatggttta | 103860 |
| aaatcgttca | gaccttagtg | ggctccagcc | tcccagccca | aagagctgtg | tgaccatgga | 103920 |
| caatttcccc | aagctctctg | ggctctcgtt | tgcctctctg | taaaaatgac | atatcaaggc | 103980 |
| tactgcctcc | ttagtttgca | gcacagatat | tatggcacia | acagatgggg | catggttatt | 104040 |
| ctggaagcgt | gtgaagagcg | ggatctggga | gaggctgggg | cagagcgctc | tgcagaagaa | 104100 |
| gcactcaggg | tgtctcttca | cttgggggac | atcaggagag | tgaacctgct | ccccccata | 104160 |
| ccagaagtgg | attccacagg | agccagttag | gctgaagggt | caggccttcc | tggcaggggc | 104220 |
| ctgagagggg | cagcagttgt | tccacagggt | cacatgttct | ggtcacaattt | gcaaaagggt | 104280 |
| ttcttttttg | ttcttttttt | tttttttagg | tttttagagg | ccttcagaaa | gcttcaggac | 104340 |
| ccacaaactc | tggaccattt | tctgcctggt | gggggtgggg | gtggccacaga | tcatccaggg | 104400 |
| agggagggaa | agagggagggt | gggggtggga | aaagtgaaat | gacttccatg | tgtgccccgt | 104460 |
| cacgagatcc | agatgtccaa | acccagtgac | cttctctctg | ccacttgagg | ggcagggggg | 104520 |
| gcaggggcct | ataggagtag | tgacttggtg | gttctggggg | ccccagcaaa | actagaagct | 104580 |
| gtaatgtagg | gagagacaaa | agggctggga | ggctcagggc | ccctgtggag | ggcggggaga | 104640 |
| catggcactg | accggctcct | ccaggctgac | gggtgcccag | gggtgtccat | ccagagccca | 104700 |
| gtgcgggggtg | actggctgcc | cagggatatg | tctcgagta | aagacagagc | acaggggtgag | 104760 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|--------------|--------|
| ggggaccta | ggaacacagg | ggcatgggac | aaagcagagg | gaggggggta | gaggacatcc | 104820 |
| ccagggaggg | actggaggcc | ttttggggca | gaacttcacct | tcaacacggcg | tgggtctcagc | 104880 |
| ctgggagaagg | aggggacgcc | gtggggcatcc | ttggatctga | ggagctatca | aggaggagga | 104940 |
| aaagagaagg | ctggaaagg | acagctcagc | tggggacacg | ggagtcocct | gaactttgtc | 105000 |
| ggggggcagg | cttgggctcg | gcgatcacaa | ggaagaggcc | aaggcccgca | gtgcagaggg | 105060 |
| gaaggaagga | gcggggcagc | cttagggatt | tttagatggg | cagcagatgc | ctttaggggtg | 105120 |
| agagatgtac | gaagagagga | caactgtgcc | cccccatca | cttgagaaaa | acaaacagcca | 105180 |
| gatgttgcoct | tgcgaggtcc | accttgccca | gagctccctc | ggggactctg | tccctggctggc | 105240 |
| agggttttgg | taccctggcc | cagaaggccc | ctctcatctc | cttcaagggg | aggggacgct | 105300 |
| tccggacgga | gccttggtgc | tccctggccg | gggtgtgcta | agggggctct | aggaggaatc | 105360 |
| ccagagccaa | gcattactca | gagggcgcc | ggaatgttcc | cctgggaatgc | tccagccccc | 105420 |
| tccactggcc | ccaaccactc | tcacaggccc | gccttgccag | agccaggccc | caggcaccca | 105480 |
| gagcctgcag | cagccctcct | tcccccgta | cccagtccca | gctcccagaa | cagacagcct | 105540 |
| ccccctccca | cgcagccctg | gcctcagtc | tgctgggctg | atggctgcct | gtggaagtga | 105600 |
| ctcagctcct | gctaggccac | cccaactcct | ttttctcct | ccacctctc | tcccagacta | 105660 |
| caaacatcaa | agacccttcc | tccaagaagc | cctccttgat | tggatgagtg | aattgcccac | 105720 |
| aggcagatga | gggcccagag | gagcttgcca | ccttggaag | gagggctagag | gggcccagtg | 105780 |
| aggggagggt | ctgagtgat | gtgggggagg | ggaaggagg | gaggtctctc | agcccagaga | 105840 |
| gcacttaact | gagagtagag | aaccagctt | tgctgctcct | aggcctctaa | gggtttgggg | 105900 |
| aagaggtagg | gtggcccgcc | gcacaggtgt | gggtgggtg | cagtggtggt | tgtgggtgct | 105960 |
| gtccacatgg | ccttgccgcg | acgtgtggc | cacgggcacc | ctgaccccaa | tgaggggagag | 106020 |
| agggggcagag | ctggagctgg | agctggagct | ccggtgaccg | ggatgaatggg | ggctggaacct | 106080 |
| gagggagcca | ggctggatatt | gggcacatag | acgcccctct | cccagggttc | ccatcacctc | 106140 |
| ccctgacccc | aggatagggc | tcagagggga | gggagcagtg | gaccgcctgg | ggccctcccc | 106200 |
| tggggccaga | acagaccagg | ccccctgtacc | tgtttggtcc | ccacacagtg | ctgtgggaagc | 106260 |
| caccgcccag | tctgcacagc | accgcccagc | ccctccctgg | ctccctccct | ttgctctccc | 106320 |
| tgttcccgcc | caggcacttg | ctgtgcagga | ctggctaata | ctcccaccog | cttgcagagg | 106380 |
| ttgttccagc | cccattctaa | catctttgtt | tggaggggtt | accccagagga | gacagctgca | 106440 |
| gtctttccag | agcactgtcta | aacagacacc | tttatctctg | agaggccctt | ctctatctca | 106500 |
| ccaaaacagg | caacaataata | aacaacatac | acactgcctc | gctgcctgag | gagggaggagc | 106560 |
| gaggggtgag | caggggtgag | gttctgcagc | ctgagcagca | agcaggagca | agcaggagct | 106620 |
| ctgggggact | cttgggcatg | ggggcttctc | agaggatgga | gccccgctga | gtcctaagggg | 106680 |
| gtggagagagc | aggagcgggt | cacacgggtg | cctgcggatg | gaagctgggt | gtgagagcga | 106740 |
| gaatccaggg | agagggggct | acggctatgg | gctgggggct | gggggctggg | ctgtccccga | 106800 |
| gggggagggga | gccatgccct | ctgctttgcc | agcggagtg | cagccgggca | gtgtgggcaa | 106860 |
| gtccggggccc | ggggccagcc | caagcacact | tgagcgtccc | tgggcaggtc | ccacggagagc | 106920 |
| ccccccaaag | agtccccacg | cctcgaccta | ctggccgtat | ggcgccgggg | ccgtgagacc | 106980 |
| ctccgcgcgc | tgacccgagc | tctgagcaga | acccatcccc | gccaccacca | ccgcgcctag | 107040 |
| cctgcccctc | agggcgccac | ccgcccgcgt | cctcaccttg | aagcaccocg | gcgcctggca | 107100 |
| ctggccagag | cagcagcagt | agtagcagca | gcagcaacgg | ggtcccccca | gctctccggg | 107160 |
| gcctccagcc | catagctgtg | agctcctcgg | cctctaggca | gcggctcgca | actccggctc | 107220 |
| gcgccaggct | ggatttcggc | cgaccctg | ccgggtgcagc | ctcaggccgc | cgccttcgga | 107280 |
| ccttcccggc | cccacctccc | accgcccggc | ctcgctcccg | cctcccctcc | ccgccaaacc | 107340 |
| cgctcggagc | ctggccaggg | gccccgacgg | cgcgcgccat | gggggagccg | ggctcgcaact | 107400 |
| cccggaccgc | cgcocctcga | gggggtggag | ctgggcggag | gagggaaatcc | gtgcggcccc | 107460 |
| tcggatgacc | ggcccgagcc | gtccctcccc | gtcgtctca | gagggcctct | actcctgaga | 107520 |
| ggagggagaga | accgctggga | aggttcttgg | aggacccggg | cgtgttggga | tgaggcgggtg | 107580 |
| ggcaaaaggc | gcctctcgct | gctgaagttg | gccccaggag | cgcgatcttc | cgtgtctctc | 107640 |
| tggggccgat | ctctgtcccc | tccttgctac | ccgtcctgcc | ccgagggtgc | cctggcggag | 107700 |
| gttgagctgg | gtcatccacc | tgcaactggg | gcccccaagg | atagggaaggt | tcaggcaacc | 107760 |
| ggctgcgcgt | gtcttggggg | cttcaattgt | gggcaaaagg | gatgcagcag | acggagacaa | 107820 |
| cctttctctc | ctggcgctgg | ccagaggcca | gaattgcata | aaagctgcag | actcccagtc | 107880 |
| ctgggagacc | ctttcgccct | cagtaacatc | tgtttcatgt | ttttaaacttt | tgttttctca | 107940 |
| ctcgggtcaa | atttggatga | gatgttaact | tttttttttt | tttttttttt | gagatggagt | 108000 |
| ctccctctgt | cgccaggctg | gagtgacgcg | gcgcgatctt | ggctcactgc | aaacctccag | 108060 |
| tccctgggtc | aagcgattct | cctcccgagt | agctgggagt | agctgggagt | acaggcgccg | 108120 |
| gtaccacccc | ccagctaact | tttgtatttt | tagcagagac | gaggtttcac | cattttggcc | 108180 |
| aggatgggtc | caatctcctg | atctcgtgat | ccaccgcct | cggcctctca | aagcgctggg | 108240 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| attacagcca | tgagccaccg | cgcccgcccg | gagatgttaa | cttttaagca | aatctttttt | 108300 |
| tttttttttt | tttttgagac | agagttttct | tcttgttacc | cagactggag | tgaatgggca | 108360 |
| tgactctggc | tcactgcac | ctctgcctcc | cagattcaag | tgattctctc | gcctcagcct | 108420 |
| cccgagtagc | tggcattaca | ggcattcgcc | accacgcctg | gctaattttg | tatttttagt | 108480 |
| agagatgggg | tttctccatg | ttggctcagc | tggtctcgaa | ctcccgaact | caggtgatct | 108540 |
| gcccgcctcg | gcctcccaaa | gcgctgggaat | tacaggcgctg | agacaccgca | ccagcctac | 108600 |
| ttttaagtaa | atctatttgt | ttttgagaat | ttggaatgta | gtaatttggt | tagtgaaagt | 108660 |
| tcgagcagtg | agagaaacct | acattcacat | atctcaaaat | caaaaagtac | agaagcata | 108720 |
| gggaaaagtc | tccgtgtctc | tagccctcct | caccaacagg | aaaccaatat | gattagtttc | 108780 |
| tttcataggc | tttttagatta | ttttttcaca | ctcaagacaa | tacagacata | ttttttctc | 108840 |
| ttattaaagt | ttttctgcac | tttgattttc | tttttttttt | tggtcgctta | atacacctta | 108900 |
| gatacagtg | cgttttagagg | gtccttgttg | ttottatgat | tattatttag | agacagggtc | 108960 |
| tactctgtc | accacgccta | gaggacagtg | gcctgatcat | gcctcattgc | agccttgaaa | 109020 |
| tctctgggctc | aaggtatcct | cccacctcag | cctcctgagt | agctggaaat | acaggcacac | 109080 |
| ggcaccaggc | ccagctaaaa | tttttaattt | ttctgtagac | agggggctct | actttgtttc | 109140 |
| ccaggcttgt | ctcaaaactcc | tggtcttggt | caggcgagtg | gtctcatgcc | tgtaatccca | 109200 |
| gcaactttggg | aggccgaggg | gggcagatca | ctggagggtca | ggagttcaag | accagcttgg | 109260 |
| ccaacatggt | gaaaccccat | ctctactaaa | aatacaaaaa | ttagccgggc | atgggtgtga | 109320 |
| gcgcctgtag | ttccagctac | ttgggaggct | gaggcaggaa | aatcgcttga | actcagaagg | 109380 |
| tgagggttgc | agcgagccga | gatcatgcca | ttgcaactca | gcctggggcaa | caagagcgaa | 109440 |
| actccgtctc | aaaaaaataa | aaaaagaactc | ctgatcttaa | gtgatcctcc | gtgatcctcc | 109500 |
| tgctcagctc | tctcaaatcg | ctggaaattac | aggagtgagt | caccacagct | gtccagctac | 109560 |
| gagattatta | cttattatta | ttttcaaatc | ttttcaaatc | aaactcata | aggtataatt | 109620 |
| tacacacaa | aaaatgcact | tatttttaagt | ggccagtaag | atgagtttct | ataagtgtat | 109680 |
| ataactacat | aagcatcact | ataatgcaga | cacattccct | cactcacaga | aagagccctg | 109740 |
| tgccctttcca | gccaaacttc | cccaactcca | accocagaca | gccactgatc | tgttgttctc | 109800 |
| tgctctataga | taagttttgc | ctgttctaga | atttcatata | aatggaatca | tgcggcatgc | 109860 |
| actcttttgt | gtctgggttc | cttccctctt | tccgatgttt | ttgagattca | tttacactat | 109920 |
| tttgcatatc | aatagtttgt | tccctctgat | tgctggaatg | tttgcgggtg | tttcaggggaa | 109980 |
| ccacagtttc | tctaactcac | agtgcacat | agggttattt | tccagttagg | ggctcttata | 110040 |
| attggaacta | tatttgcaca | gagagagaga | gaggaagaaa | gagggagaga | gatattttat | 110100 |
| atagcaattg | gctcacttga | ttatggagtc | caaaaagttc | ccgaatctgc | ctatctcaag | 110160 |
| ctggagaaag | aggaagaaca | gtggctgtat | tcagtttgag | ttcaaaaggc | tgagaaccag | 110220 |
| gagcaccagt | atggaagttg | ctcgagctca | gaacaaagtg | gggacaggaa | agcagagcag | 110280 |
| cacccacag | cagccctcca | gcgacacctc | ttcagtaaa | caaggctgaa | cacagagggg | 110340 |
| ctggcttcag | tgtggatgtc | aggtacagaa | ggcagctcga | ggagctactc | tggcgttctt | 110400 |
| gcttactggt | attcttacct | cgaactggcc | aactcctact | taaaactcgag | gccatggctt | 110460 |
| taatgtcctg | tcattcagag | gctgtccctt | acccaaagcc | agggttagcat | cccctgactg | 110520 |
| acactctctc | ctgcaaacag | tttcagaagg | ccctgtagt | gtccactctc | ctgtctctct | 110580 |
| cccaagcttc | ctgagctcca | tgttgtctgg | gaatatgtgt | gttgctcact | tcctagcaca | 110640 |
| gtcagtgcta | ataactgact | gtagagggga | cacagtcgaa | aagccacatg | gggatcagag | 110700 |
| tcatctttac | acagttgaca | cctcccaaac | ccagatgagc | tgtgtccaag | tgcagggtcag | 110760 |
| aggaatttttc | tccggaagtc | ggtttattta | ctattttagg | ctattttagg | ttgcaggggg | 110820 |
| ggagtagagg | ccatcaaaac | aaagcttagg | aagagggtac | ctaggaatga | ccgagcagct | 110880 |
| ccggggggcg | ctgacagcac | ctgggaagaa | tggcttctcc | actggcttgt | tgccgtcacc | 110940 |
| ctccagagg | gcatcaggaa | atgtcctggg | aaccaggcaa | accagtgagc | attaaccctt | 111000 |
| agaagtgcct | ggcatgggtg | acacccacca | ctgttaaaaca | cgactctctc | caaggagtga | 111060 |
| cgcagacag | gatgtctgag | ggagcactc | cctcagagat | cttcagagat | cgcagggtg | 111120 |
| gcacctgggt | acgacaggct | gatgcttggg | tgccccagaa | aaggctcatgt | gtgtgaatgg | 111180 |
| gggccccaaa | gccaacgcct | catccctgac | agcctgggtc | attttagagg | gaactttttg | 111240 |
| tcccttggca | aggtgggttg | aatttcagtg | tcatagggca | aggttatttt | agcttttaata | 111300 |
| gatatgttca | aacagttttc | caaagtcat | gtcacacact | tgtgatttcta | cttatgtaaa | 111360 |
| gtttaaaaac | agggcaaatca | aatctatggg | gttcgaagtc | aagacagtag | ttaaccttgt | 111420 |
| ggggggtgca | actggtacag | ggtgtacagg | gggactgtag | gatggtctat | ttcttgatct | 111480 |
| gggtgtgttc | gcttttgtaa | aagtccttga | gttgcattta | taatgtgtga | acttttctgt | 111540 |
| atgttacaact | tttaattgaat | gtacaaaag | tctcaggagg | cttcagacga | cttggaaggc | 111600 |
| acacaactaa | cccctctgag | agcctccaat | ccaagatgga | catatgtccc | cttggaagta | 111660 |
| tcgacaagca | ggtgaagact | cctaagccgg | atatcccaa | atccccccag | tagcccgagc | 111720 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ttcagcagct | gcttatggtc | ctccctacac | cctctcttcc | ccagacagcc | cccaaacatc | 111780 |
| tggtgcatt | tgacttgctc | tctccctgtc | ccacctctgg | atttagtcca | tgtttctccac | 111840 |
| cctccccact | gtcagcaatg | tagacaagac | aaacgcttag | tccacgtgccc | cacctactgc | 111900 |
| gtgcccagca | cggggctggg | catttgggtg | ggcaaatgtg | agcaacacac | gaagcctcaa | 111960 |
| ggagcagaaa | gggacacaaa | tcacttcagc | gtaaggtaat | ttgtgataaa | tgtcatgtaa | 112020 |
| cttcgagccc | ctggccccct | cctacagatg | gtgtctaaga | ataaacccca | ctaactatgtg | 112080 |
| actcctctgt | cttgaccagc | ctgtttgggt | tgcaagaaa | agactcactc | cagtttgcgtc | 112140 |
| ataggatgga | gttttatctg | gaggacattc | tggacgggct | ccagacaaga | gtctggcaat | 112200 |
| ggagcatgaa | aatgaatgag | cctggaatga | gggagggtgc | ggcctcggct | acacaaaagt | 112260 |
| caccgggccc | ccaactgcct | cccagcggtg | tagctcctgt | ggcaactccc | cacttctctc | 112320 |
| tctgtttttc | aacccaaatc | ctagagcaga | gggctcttcg | ctcctcacac | catctcccac | 112380 |
| aggctgcagg | gggagtcact | agcccactca | agagctctct | cctgttggtc | cctgattcgc | 112440 |
| aggggcacgt | cattccttct | tggttaactca | ttctctttaa | cagcccgaca | cagggtctcc | 112500 |
| agggaaaagga | caggagctgc | cacagctcatg | aacagagatg | catctggagc | aggaataaagg | 112560 |
| atgctgacaa | ctctgtgttc | tgccctccac | cttgttaagt | gccagtaagg | tagtaccocag | 112620 |
| tccctggggg | tagggggggg | ggtggtggga | atgaggggcat | ctctctcttc | agggggccaaa | 112680 |
| tctggcacaa | ggatgtctgc | ttaggcaact | ccactgctctg | gcattgctctc | tccctaggaa | 112740 |
| aaacacaaac | ctttttttat | tctcagctct | tagtgtgaga | gtgactcaact | cttcagggaa | 112800 |
| gccacacaac | agaatggcca | ggggcacacc | ctggccctca | gtagatgggc | actgagacaa | 112860 |
| aaactggcct | gctggtggca | ttcccaacaa | tgtcaaaagt | ctcagggatg | agctcacagt | 112920 |
| tgggagtcct | ttgggaagtc | aattccaaat | gtcctctgga | ctcaaaaata | aaactacatt | 112980 |
| tccagtgctc | ccttgtagct | tgcagtagcc | atgtgacaa | actccagcca | atggggaagtg | 113040 |
| agtggaaaa | tcttgtcag | ctcttgagtt | gcgtccacca | tgggcatggg | tgtgcttcta | 113100 |
| cctccactct | tctctcttct | ggctggagg | catgtggaca | aagcggggcc | atggtgagcc | 113160 |
| tcacaatcaa | aggcatcatt | ttagggatag | agaaaaaggaa | gatagaagg | tccctgagccc | 113220 |
| caacatcata | aacctcgtta | acagccagac | tagcgtggga | gagaaaaata | aaattctatc | 113280 |
| atgttgaagt | cactgtattt | ggtctcttgt | tagagcaggc | tgaactgatac | tctgttgaat | 113340 |
| acagaatgcc | ttggtgagcc | tctgaggatg | caggatgctg | caatcccaac | aaagtccagg | 113400 |
| gtgacttctg | aggagaagg | tgtgaaagg | cagggtcttc | tttgaggggag | aaatggagag | 113460 |
| aaaggaggag | atccagagaa | gagagcagca | gatcagctca | gtgctcttca | gtgtcctcca | 113520 |
| gggtcaacac | cctcctgggc | agactcttca | ccaccattt | ttgggggggc | aaagggtctg | 113580 |
| ggggccaggtc | ataaaaaagtc | agatgtcaca | gagctgtttt | cgtagaggcg | ggcaggactc | 113640 |
| aacactgcct | cattcacatt | cataggtctgg | agtcatcaca | gagttctggcc | acttctctct | 113700 |
| ccggagggct | ggcaacacca | ctggtcagcc | cagggggtggg | gcacaggttg | aggtctcaca | 113760 |
| tgtggctgca | tcaggatcaa | tgagcttccc | acagagaaaa | cctcgtcac | aggaggctc | 113820 |
| tgggagccag | ggcaggcccc | caggccccca | gcctcagggtg | gggtcccaag | aggattgggtg | 113880 |
| ggagtggtgca | tggctggaag | tgtgacttgg | agccctggcc | taccagggtga | atcagctgca | 113940 |
| gaagggacac | accacagggg | aggagaaact | gactttcact | tctgccagg | tgcagggcag | 114000 |
| gcttgggagt | ggggaagagg | cctcagaaca | aagggtggggg | agtcagaaca | tacgtgtctt | 114060 |
| ggaaatggag | gggatgctgc | caagtccaga | aagtctcccc | tagtctctcta | agccagcgcc | 114120 |
| cagctgacac | agggattccc | tctgccccga | gcagaaacagg | gcctctctat | ctcctgaggg | 114180 |
| gctcctgggt | ccctgcacac | tctccccaac | ctccccctgg | tccaggccac | catcctctaa | 114240 |
| atgacagcag | tctccagatg | ccagttggca | ctgtgggtta | gctgggtatt | atatctgccc | 114300 |
| gtgtgtgttg | aatggatacc | tgcgtgggtc | ctttctgaaa | agccattctcc | accgaattct | 114360 |
| gcgccatgct | ctgcttacaa | ttctgcaagg | ccggagtgg | actcagagag | ctgtagcag | 114420 |
| cctctccctc | cccacactcc | tgccccatt | cctctcccg | tccccacagg | ctggctccag | 114480 |
| accacaagag | cgtgggaacc | tcactggagag | tgggtgcttc | ctctgtcttc | ctccctagcc | 114540 |
| cttgtcttag | cacagccagg | gcagagagga | ggaaaagaagg | aaactacagc | aaggaggatt | 114600 |
| tagggacaat | tctagaagg | atttccaatt | tggaggggca | gtctgggaag | taggcagcct | 114660 |
| tggaacttaa | taggaggtcc | taaaaggact | gggggaattt | ggggagaaga | cgagggtggg | 114720 |
| gtccttacc | caatctctg | ggcactccc | agctccagac | ccactctcag | gtgtagcag | 114780 |
| cccccaggcc | caacagcagg | agcaggaggc | ccaccagcag | tcccaggacc | cagagcagct | 114840 |
| gctggaactg | atgcaggcgg | tgcagggtcg | gaacacagag | cgggactcag | agcagccaca | 114900 |
| gctgcaggcc | cccagtgctt | ccttccaagg | ggacccaccc | aagatgacac | cttctaactt | 114960 |
| ttgctttatc | tgtgaagctt | tccacaaga | ggtgccccga | gcaatcagtt | tagagctgac | 115020 |
| aggcaaatca | tgctgcagca | ggagggatc | aggggaaggaa | gccttctggg | ctaagagctg | 115080 |
| ggcttttgtg | cagagtgccc | tgggagatgg | agccccctcc | cctcactctc | gaccccaag | 115140 |
| taggtggagg | tagaggcaga | gcccaggaca | tttgaggcag | aacagatgta | ttccccctgg | 115200 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| tcgctggggcc | tcagcgccctc | gatgtccacc | ctcagggcat | tggggggaagc | caggacatgg | 115260 |
| atgtgtggct | ctgcaggagc | acccctggggc | tgactggagg | ccaccagtcg | actgccaaag | 115320 |
| tggagagtca | ggctggcgag | cggctcgctg | tccactcggc | aatccaggat | gccccggagg | 115380 |
| ccaccctcag | gctccacgaa | gaccatcatg | gtgggcgctc | tgggagggtc | tgtggggagg | 115440 |
| aagggagtgt | ggtgatgtca | gaaaatgacc | acaaaattct | ccctccctgt | acccatgcac | 115500 |
| agtgaacttg | ctgtcctctc | cattaaagag | cagaacctat | ttcctcaact | cttgaatctg | 115560 |
| tgactcagtt | tgaccaattg | aagaaagaa | tgatgttggt | caacttcaga | gccagacctc | 115620 |
| aagaggcctt | gtagcttctg | ctctacttgt | tggaaacacaa | caatgtggga | agcccaggcc | 115680 |
| agcctgtctg | gcacatgtgg | cccaactggc | agccaggagc | gtgcagtcat | ctgagaccac | 115740 |
| tgggcaactg | cagccacatg | aggtgtgaac | aacagaagaa | ctccccagct | gagcccagcc | 115800 |
| cacagaattg | agcagataac | atgacttttg | tttgaagcag | taagttttgt | tgtggctcaat | 115860 |
| agcaatagct | gactgataca | gtgtaggcct | tggtaagggc | tggagtggga | gaccaagact | 115920 |
| gatgagggtc | tatcgttgtc | cagacacact | cagcacgctt | tattttctttt | ggtttttttct | 115980 |
| cttctttttc | tttctttttt | tttttttttt | ttttgagaca | gagttctcagt | ctgtcaccca | 116040 |
| ggctggagtg | cgttctgggtg | actctggctc | actgcaacct | ccacctccca | ggctcaagtg | 116100 |
| agtctcctgc | ctcagcctcc | tgaatagctg | ggattacagg | tgtgtgcctt | catgccgtgc | 116160 |
| taattcttgt | attttttagta | gcagccggagt | tctgccatat | tgaccaggct | agttcttgaa | 116220 |
| tctgcacctc | aagtgtatcg | ccacctcag | ctctcctcag | tgtctgggatt | acaggcatcaa | 116280 |
| gcccttgca | cagccattta | tttcttatac | aacttgcagg | atgaggcgaa | tgatgtattc | 116340 |
| cccattttac | aaggtcaccc | aaccagacag | cgggagagcc | aggattccaa | cccaggactc | 116400 |
| aatatccgtg | aagccttcac | tcttggcgaga | cagaagagga | ggaagaaactg | aggtgggcgt | 116460 |
| gtgagatgtg | gggagactaa | ggcggaggag | ggggttcaca | ctcacagagc | acacggagca | 116520 |
| tgaactggagc | agaggcgaga | gccccagtg | ggagctcagc | caggcagtg | tacatcccag | 116580 |
| cttgagcacg | agccacgtgg | gtgaaggcga | gagtgggcac | aggctcccg | ggagccgcgc | 116640 |
| ggctattcca | gaaccatgca | aagggtggagt | tggccacagg | ccaggggcca | ccagggagcg | 116700 |
| ggcagctgag | gttcaggcca | gcgcctcag | gcacgtccag | gccaggctct | gccaacctgc | 116760 |
| gtgcacctgc | gggcggagga | tagagagatg | attggggatc | tgtaggcctt | gggggcttga | 116820 |
| gcctctgggt | gggacctctg | aggggacctc | gcacatttgt | ggaagagtctc | agttctgactt | 116880 |
| ggtatagggc | tttccaaggt | ggaacctcag | ccctccag | ggagagctgtg | ccccctcac | 116940 |
| cagataggcc | ttctcaaggc | agagctttct | cagaccagac | agcccgcctc | agtgggagct | 117000 |
| ccagcccctc | tctgtgactt | ggacctgaga | gcactcagagc | ccctctctct | ccctctctct | 117060 |
| ccactcacct | tctactgcga | agcccccag | ggtgctgatt | gagcccagca | agttttgggc | 117120 |
| tgtgcaaaac | tagtgtgtcat | caacttgcagg | cacatcttgc | acctgcagcc | tgtagggcgtt | 117180 |
| tcggggccacc | tggacatggc | ctgtcctga | tgccaaagctg | tggaccctgc | tgctcgtggc | 117240 |
| cagcaccttg | ccatcatgag | atgagggccag | ctcagcaggt | ggctcactgt | ccacagtgca | 117300 |
| ctgtatcaca | gccatggatc | tggccctgga | gtcccgggaag | gaggacagga | cagcgtctgt | 117360 |
| aggggcactc | gtgggcaggc | agggcacaga | tggggacctt | gcttaggcac | ctgtacttgc | 117420 |
| tcatltgccta | gctgcctggg | gttggtcagg | ctgaggcctc | tggaccaatg | gccacttctc | 117480 |
| agaagtgcga | ccttcccagg | agtctctgtg | gcagagtgtc | ccaggcaggg | cttataggag | 117540 |
| tgtctctttc | tctagcacct | atgtcctctt | tcccacactg | ccatttctctg | ggccacttgc | 117600 |
| agtcttctgc | ccactcccac | cagagcccag | aagcacctct | caccacctct | ctcctgggca | 117660 |
| gccttggcca | aggaacctct | gctcacagag | gacttgcagg | gcagcaggac | gggagctggc | 117720 |
| ggttgccctgg | gctcctggg | gctgccaaga | gtaggcgctc | gcactagccc | cgctggccaac | 117780 |
| cagggaatgag | agtgaggcag | ctggaccctc | ctgcagccaa | cgacgcttgt | ggtaccaagt | 117840 |
| atagagtgtg | ggtgcgtggg | cagcagggtc | cgcacaggtc | acttgtatgg | gggcaacctc | 117900 |
| aggcacggca | gcctccggag | ccaggatcac | ccgcacacct | gtgccaagga | ggccaggatc | 117960 |
| agccggggccc | agccgggacac | cacagttggc | ttccatccca | gtgggcttgg | cctaggccct | 118020 |
| gccttacctc | ccagcccgag | ctccagggac | gtgttggcct | ggcccagagg | gctcgggca | 118080 |
| gagcagctgt | agaaacctc | atccctgggc | tgtggccctc | gcagctccag | gcgcagggtg | 118140 |
| ttgggggacag | aggtctgtgt | cgaggaggcc | aagaggcgac | cggcgtggct | gagggccagc | 118200 |
| tggggcgggc | gtcggtgtgc | agtaggcac | agtagccagg | ccagctgcc | gcatggctc | 118260 |
| tccaggaggt | aggtcaggcg | caggttgcgg | ggcgcgtctg | cagggcatga | gaggcttaca | 118320 |
| cggagatcaca | ggcggcagcc | tgcaggaggc | cagtttgcag | gtggcattca | aactcaggag | 118380 |
| ggccctggct | ccccacctc | atggctccca | accacctcag | ctgaggcaact | gtctctctgc | 118440 |
| ccagacagga | ctcacccag | cgccccctac | tcttaatgct | ctctgcccag | tgtctcacta | 118500 |
| gctactgggt | ccactggatc | cctgtcgga | cctgtcgggt | cccgccatg | cggtgatccc | 118560 |
| tgtgggtctc | tgtcattcct | gtgagttcca | ccctgcattc | agccagactc | acagaggagc | 118620 |
| tccaaggtga | taggtctgga | gaggcggggt | gcccgaaccag | ggggggccac | accgcagcgg | 118680 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| taggaggtgg | catccctgac | tgtgacgttg | ggcaggggga | tggagtgggc | atccaggcgc | 118740 |
| tgctgcccat | cctggtacca | tgtgtagggtg | agctgggccc | gggtgagtgg | ccacacaagg | 118800 |
| caggtcaggt | tcaccagctg | ccccctccgc | acggtagccc | cggggccacac | ctgcacattc | 118860 |
| acagctgggg | agagggaggg | cacaggacac | cagtgagggt | cttgaggctg | tgtacagcag | 118920 |
| gccacctgtc | tccatgtggg | tgagctactc | ctactgctgg | ggagccctct | ggcctttggg | 118980 |
| agccttgggt | ggccccacta | taaatgtggc | ttagaactca | agcttgggaa | tagcccaact | 119040 |
| gctcctgagt | ggctccaaat | agagttccac | agagtcctgg | aaagccctga | gccccctccc | 119100 |
| gaccctggct | gggggcaaat | ggggagacca | cagacctccc | ctccccctgt | gcactttctc | 119160 |
| ctgaaattcc | ctgtttagtt | ttatccactg | ggcatacttg | taaaatttgg | tgttttataa | 119220 |
| aggtttcctc | aagaaaactc | agtctggagt | ccaccatcgt | taggaaaggg | gggaatgggg | 119280 |
| actattcccg | tgcccccagg | cttctagaac | cctgtccccc | ctcctctcct | cttttaaagg | 119340 |
| acacacacac | acacacacac | acacacacac | acacacacac | acactgacct | tgagcgtcga | 119400 |
| agtacgtga | ggccgaggcc | tggcccaggg | tgttggaggc | ctcacagata | tagacaccct | 119460 |
| catcctccag | catagccccc | tggatctcca | gacgcagtg | gttggggggc | acagccacat | 119520 |
| gcagcctggg | agagctgcct | tcgggtcccc | ccacaccttg | tagggctggag | gccacaaggg | 119580 |
| gatccccgtg | gagcagccgc | agctgggccc | gagggctcact | gtccacacgg | cacaggagga | 119640 |
| ggccccagtc | tccaggacct | gtgtccatca | gggtagttag | tgtgcagtgg | cgtggggcat | 119700 |
| ctacacagg | tggggagtgg | tcaggaccga | gccagggggg | tccctcatcc | agggctctgt | 119760 |
| catgtgacac | agcccaggac | tgggggactg | catttccttc | ccacacacct | gatgggacaa | 119820 |
| aagtccctac | aggacacgtg | gaggctgatg | ggtgcagcta | ggctcgtgg | ggctgagcct | 119880 |
| ggggccctgg | cttggccaatg | ataggcccca | gcttgtgtca | aagtttatgg | tgcaaaagcgg | 119940 |
| agcgtggccg | aggtcgactc | ctggaggggc | tggccatccc | gataccaacg | atatgaggct | 120000 |
| ccctctggga | ctcctgtggg | tacctggcag | ctcaggacca | cagcctggcc | ctcctggagc | 120060 |
| tcaggtgatg | gtgacacctg | gaccaggctc | cctgcagggg | aaaaccaaga | gcaggttagg | 120120 |
| gctctccacc | acacctctca | cagttctggga | ccatgtgcgt | gtccacctag | gactacacaa | 120180 |
| cccaaaccca | tgtgctggag | ccgagccccc | ttccagacca | ccgcaggggc | ccactcagcc | 120240 |
| ctatgccttg | gccctcctg | ctccccactc | tccctgcacg | gtcccatcct | ggcatttacc | 120300 |
| tcccaccacc | tgggtgggatg | gttttactgt | gtcagtcata | tccgcctac | caaagcgtaa | 120360 |
| gccttatggg | cactggactt | gtgtgacctg | taacccagtg | ctcagggcat | gtgcctgcct | 120420 |
| gtgttagctc | ctgacaaatg | ttggttcaac | ccataaagta | ctgaaaaggg | agggatttag | 120480 |
| atcatccag | ggccatttgt | ctaactagga | tctggactca | gcattgggtga | ccgagggctt | 120540 |
| ggaagggaa | gttagcagct | atgttatctt | gagcgctctc | gccagggccg | ccctgcattc | 120600 |
| agactccagg | ccacaagagc | acaggacctg | gggaccaggc | aaaggcgacg | tcctcagtg | 120660 |
| ccctcgggtg | tgaagcagga | gaaacccagg | ttgcagggg | tgagatggag | ggacagctgg | 120720 |
| aatgctaagc | aagaacacag | accatgcctg | agaccacagc | tagcagggtc | acgcaactgg | 120780 |
| ggaaaagctc | ctaaaaactc | aaccttcaat | ttcttcctgc | gaaataggta | tgctaagtaa | 120840 |
| gaaaagatac | acagaaccaa | ggctcaaaat | aaatgtccca | aaattggctg | tgtcgttcac | 120900 |
| cgtggtggaa | gaaacactca | tgttcccttg | accttcgcct | tatgggcatt | atctttcctt | 120960 |
| tttcctgggt | tatttgttta | ggggagaattt | ttttttttta | agacagagtc | tcactctgtc | 121020 |
| atccaggcta | gagtgaagta | caattttggc | tcactgcaac | ctccatctcc | caggttcaaa | 121080 |
| cgatctgcct | gccttagcct | cccagtagaac | tgggattaca | aatgccacc | accacacctc | 121140 |
| gctaattttt | gtatttttag | tagagtgttg | gccaggctgg | tcttgaactc | ctgacctcaa | 121200 |
| tgatctggcc | cgctccagcc | tcccagaagt | ctgggattac | agggcatgagc | catgggtccc | 121260 |
| ggcctgaggg | aggattttta | tagagactat | atataatata | acacacacac | acacacacac | 121320 |
| acacacacac | acacacacac | acgtatgtgt | atataaataa | agaatatgaa | tatatattgc | 121380 |
| ataagaatac | acataatgta | tatttgtatg | tgtatatact | catacataaa | catatatgca | 121440 |
| taataataat | atacataatg | atgagaatgt | ttactatgta | acaaaattat | atgatcgcaa | 121500 |
| aacaagttat | ttcagaagtg | gtgcccacca | agcaagccac | tttgtgaatt | gcatctgccc | 121560 |
| cctagtggtc | gctgcagtta | ttacaggtga | aaatatctag | ctggaggcaa | agggaggcct | 121620 |
| tctgtgggtc | ggctggaaaac | tgcccttacc | aggactgaaa | aagacatatt | tctgcacaaa | 121680 |
| attcacagca | cccagcccaa | agttctgacca | gagtagtccc | cagggcatca | gagtcacca | 121740 |
| tatcttcaga | agactgacac | tcacctcggg | cctggaagaa | gagtgagggt | ttggatgatc | 121800 |
| caagtaacatt | tgtggcctca | cagcggtagc | tgccagagtc | cccagggccc | agttctcggg | 121860 |
| cctcacaact | cagggagttg | ctcctcagct | tagcctggaa | ccgacatagg | tatgggaact | 121920 |
| ggggaccagg | gctgggtggc | aggaggtgct | ccccatggaa | caaggccagc | aaggccaggg | 121980 |
| ggcggctgtc | cacagtcag | atgaacagag | ccatgtggcc | ctgggcccgt | cttgaggagg | 122040 |
| ctgacagctt | tggacggtcc | gggggatctg | caggaaacaga | gggagctgag | gccacccagc | 122100 |
| ctagctcaac | acattacaac | tgccacacta | tgtcccccc | ctcctgccac | acacacagcc | 122160 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| taagccatgc | tctttcctcc | ccaaacccca | gcccgccccc | tgtcttctccc | cagaccacccc | 122220 |
| ctccaccctc | caagtcccca | ggctgcccat | gccagtcacc | cacagaggac | actcaggagc | 122280 |
| acgggagtgg | agagctgggg | accagcctca | gtcaggatgc | ggcaggcgta | aaggggcagca | 122340 |
| tcagttctgg | ccaaggggcag | cagttgtcacg | gtctccaggg | gacctctggg | ccacagcacc | 122400 |
| ccatttcgga | accaggagaa | gttagcaggg | ctgccaccag | cttcccggct | cacgttgcaa | 122460 |
| gtcaagttgg | cttctgtgcc | ctctctaagt | gtgtgtgatg | gtgcaatggc | caggacagtg | 122520 |
| gctggagagc | agggcggcaca | gctttactgac | caccocccgc | cccaggagc | caggggtcca | 122580 |
| gcgcgccagc | ctggaaggag | caggaaaggag | gcccccctggg | gtgccacag | ggttggggag | 122640 |
| aaaagcaag | tagctcacag | ggcagcctag | gagaagtggg | ccctggggac | tgttggggcc | 122700 |
| ctgggcagag | aggggtgcag | tacagggaag | gaggacaagg | cagcccaact | gtcccagctg | 122760 |
| gggagtcctt | cctcagagac | cagctgtgtt | gcctatcccc | ggcctgaaca | gagatcatgt | 122820 |
| gaaggagcag | ctccgctctt | cctgaatgtg | gaggagggca | gcgaagggcc | ccactcgtg | 122880 |
| cacagagtgg | ttttgcctga | ttagatcctc | ctcggagcag | aacagtcag | agctccagcc | 122940 |
| cctgccccca | ggccaccctat | tccttgctctg | actcaccctg | gccattgaag | gtggctgagg | 123000 |
| tggaggcggt | ggccaggcca | ttgctggcct | cacagaggtta | caagccctcc | tcttccagca | 123060 |
| aagggttgtg | aatctccaca | cgcagcaagt | tgggggcttt | ggtgaccttc | atgcgtgggg | 123120 |
| aacagccccc | acaggtgctg | cagccacccc | ctgatggcag | ggaagtggcc | acaacacggt | 123180 |
| ccttgttggg | cagctgcagc | ctggccgggg | ggctgctgtc | cacacggcac | aaaaggaggc | 123240 |
| ctcgccttcc | agccccggcc | ccagcgccat | caaggtccag | cctgtgtgtg | aatgttgtgt | 123300 |
| gtcgaggggg | gtctgcaggg | aggaagaaca | tggggactca | tccccagat | gctccagggc | 123360 |
| cccacaagcg | tggctaggct | cccagaatac | actggacaaa | ggcagatccc | gaaggctgcc | 123420 |
| ccaagcagct | gtcgagactg | tgcactgcac | aagggcagcc | agaccagggt | gggtggaaat | 123480 |
| gtcgctaggc | tcatgtctcc | caagctaaag | tgtgccctga | tgccctactg | gtcttcccag | 123540 |
| aagaaggagc | accattttct | agctcagaca | aaggtgctgt | atgggctggc | tgccaccttg | 123600 |
| cataagtcct | atctgctgct | gagcagggtg | cccacctagc | aatccctggg | ggtcatctgt | 123660 |
| tctgtctcca | ccctgcaccc | cacggagcct | ccctccaaga | gtctccctgc | gtcttcccag | 123720 |
| agcttatctg | tgatctccaa | gccagctcct | gtgatactag | cgaatgcata | gaatggcctc | 123780 |
| ttcctgaagg | cttctggaga | cagttggcca | aaagagagca | atcagccaag | gcccccatgt | 123840 |
| gatctcactc | agacaccaga | aaacttttga | gaggcagctg | ccctctgaca | ggttcggggc | 123900 |
| gtggctccgg | cacacatgtg | tcaggggaaag | cctctggaat | caggcccttg | gtcttaacag | 123960 |
| gagcctcccc | gcttgccgga | tgcagagggg | ctgtctgtcc | ctgtctgtag | gggctttgcc | 124020 |
| ttggtgtcta | tgggagccca | gcacagcctg | tggggcatgt | gcacaaatag | aggagtgtgg | 124080 |
| tgtcccagtg | ctcacccatg | aggacaaagg | atggatgtgg | tgccccatgt | gagctttgag | 124140 |
| ctgggacaga | ggagcccaca | gggctggcat | tggagggtaa | agacatgggg | aggtagagaa | 124200 |
| ggccccagcc | tcatttccta | aacttcagcc | ccacgtggaa | agaccactg | tggccaggtc | 124260 |
| tgtgaccag | ccctcccgat | atagaccttg | gggcaggggca | ggacagggct | ggggggccta | 124320 |
| gaggagtggt | ggtggctggt | gggtcccggc | aggaggctgc | aacaggtggt | ggctgtccac | 124380 |
| agagcacagt | gagaacagct | ggcgaaggag | ggccactggc | actgtggcgg | tcccggggcc | 124440 |
| ggcagtggtta | tgagccggcg | tcagtgtctg | aggcccgggg | gagcaggagg | ctgctgcggg | 124500 |
| gacctcgtg | aagcagggct | ccattcaggt | accaggagaa | gcgggcatac | ggtgtggggc | 124560 |
| ttaggccgct | tctgcagctg | agtgctactg | cctgtccttc | caccacctcg | gctgcggggc | 124620 |
| tgatgaggag | acgggcggct | gcggggagag | gaagaggctg | ggaagggtcc | ctcctctcaa | 124680 |
| ccccacatgc | tgccctatat | agaaagcctt | ccagggttct | cccttcccca | cactactgta | 124740 |
| ggtagctctg | ggcttttggg | tgcatcagga | gggttttgccc | tttaatgcca | aatcagatct | 124800 |
| attggtagta | gcattgctga | gcattcagag | agattcagag | ccagagccca | ggcttatgtc | 124860 |
| caacacactgg | cttggctggg | catggtagct | catgcctgta | atcttagcac | tttgggagac | 124920 |
| tgaggcgaaa | ggatcgcttg | aggccaggag | ttccagacca | gcctagggca | catagtgaga | 124980 |
| ctcatctcca | aacaattttt | ttttttttga | gactgagctc | cactctgtct | ccaggctctg | 125040 |
| agtgcagtg | tgtgatcctg | gctcactgca | acctctgctc | cccaggttca | agcgatcttc | 125100 |
| ctgcctcagc | ctcccaagta | gctgggatta | caggtgtgcc | accacacca | gctaattgta | 125160 |
| tacatgtagt | agagataggg | tttccacatg | ttagccaggc | agatctcgaa | ctcccagctc | 125220 |
| ctggtgatcc | accgcctcca | gcctcccaaa | gtgctgggat | tacaggtgtg | agccaactgtg | 125280 |
| gcgaagcttt | tttttttttt | tttttttttt | ttttgagacg | gagctctgtc | cgctcagcca | 125340 |
| ggctggagt | cagtggcgca | atctgagctc | gctacaacct | ctgcctccca | ggttcaagca | 125400 |
| attatctctg | ctcagcctcc | ctagtagctg | ggaccacagg | tgtgtgccac | cacacccggc | 125460 |
| tattttttgt | atcttttagtg | gagatggggg | ttcaccacgt | tggccaggct | gatctcgaag | 125520 |
| tcttgacctc | aggtgatctg | cctgcctcgg | cctcccaaa | tgctggaatt | acaggcatga | 125580 |
| gccaccatgc | ctggccacaa | tttttttttt | ttaattagct | gggtgtgggt | acatggggcc | 125640 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| tagtctcagc | tattttgggag | gctgagatgg | gaggatggct | tgagcccagg | agtttgaggc | 125700 |
| tgacgtgagc | catgaacata | ccattgcaact | ccggcctggg | caacagagca | acaccctatc | 125760 |
| tcaaaaaaca | aaaagaaaaa | cctggcttga | tcaatttagct | acccatgcctt | caggaggagg | 125820 |
| gaaggacagt | gcacataccg | aaagttggaa | gaccgtactt | ttcttttttc | ttttctcttt | 125880 |
| tttttttttt | ttttgagaca | gagtctcaact | cttgtcaccc | aggctggagt | gcagtgccgc | 125940 |
| tatcttgccc | cactacaact | tccacctctt | gggttcaagc | gattctcctg | ccttagcctc | 126000 |
| ccaggtagct | gggaactacag | gaactcacca | ccatgcccag | ctaatttttg | tatttttagt | 126060 |
| agagatggga | tttccacatg | tgggccagg | tggctcogat | ctcttgaaact | cgtgatccac | 126120 |
| ccacctcggc | ctcccaaagt | cgtgggatta | caggagttag | ccaccgcgcc | cagccagaag | 126180 |
| accctacttt | tctattttggc | ttccacatc | tgactgctag | catagagcct | gctcccagag | 126240 |
| tttcataatt | aaaaaacat | gaatgcttct | gagggactct | ccaagttag | ggtcagggtg | 126300 |
| gggtcaaaaag | gaatgatgtg | acctgttgtg | tttccctttt | tccttgact | tcagggaagc | 126360 |
| tctgccgttg | ggctactgca | cagccctgt | cttttatgtg | gcgtagccag | ttagctcagt | 126420 |
| cctcggtgtg | agtcacactag | actcttagaa | ggaacagact | ggagcaggct | cctcctcagg | 126480 |
| ctcctccac | ctcctcggc | tggccagtgc | catcttacca | ttggcatgga | agtcagggtg | 126540 |
| ggaggttgca | tttccaagg | agttggtggc | tgagcacttg | tactccccac | tgtcagtttc | 126600 |
| ctccaggtct | cggatctcca | ggcgacagg | gttgggacca | gaggtaccac | tgaagcgtgg | 126660 |
| gctgtgatca | ctgtccccgg | aggtggaggc | caggatatga | cccccatgtg | acagcacagg | 126720 |
| tgtggccagg | ggctcactga | ccacagagca | gtgaaggatg | ccacacaagt | ccgcctgggt | 126780 |
| ctccaggaa | gctgtcagg | ctggagttag | aggcgggtct | gcgtggaagc | gagagggtgg | 126840 |
| cctgtcaccc | tcagacaagg | gcattccctg | gataccctga | tcaaccctg | gacctctgca | 126900 |
| ccgctttgtc | ccacactgcc | ctgcgagaag | ggggtgatcc | caaagtgctt | gagtgctgtg | 126960 |
| tgaatacaact | tttgtccttg | gctggcttgg | gtacgtcaact | ctgttgtccc | agtcagtggt | 127020 |
| caaggccacc | ctgcaagtgg | ggataaaagc | cccactcaa | agatgagaaa | actaatagag | 127080 |
| agacctgggt | agggacagca | cctcagccag | gtgaccaaa | tgagcatcat | cagcaatggg | 127140 |
| acaagctgac | acagactgcc | tcttgacagg | gcagagcagc | atgtccgcgg | ccttcccacc | 127200 |
| cagagggcat | ggctccaggc | cattcaagtc | aggcaaatgt | ggagttaggg | acctcctcac | 127260 |
| aacagcaggg | ctgcactctg | caagtttcaa | cgtcaagaat | gacaaggaaa | gactgaggaa | 127320 |
| cagtcctcaga | ctaaccagaga | atgaagagac | gcaaccgacc | aatgcaatat | gtgaactgtg | 127380 |
| attgtattct | aggaccagaaa | aaaaattggc | acaaaagaca | gtattaggtc | caactggtaaa | 127440 |
| atgtcataat | gcattataagc | ttagataaca | gtcttctatc | agccaggtgt | agtggtctcat | 127500 |
| acctgtaatc | ccagtacttt | gggaggccga | ggttaggtgga | tcacgtgagg | tcaggagttc | 127560 |
| aagacatcc | tggccaacat | gggtgaaccc | tgtctctaca | aaaacacaaa | aattagccag | 127620 |
| gcaggatggt | gggtgcctgt | actccagct | ctcaggagg | aggaagatcg | aggaagatcg | 127680 |
| cttgagccc | agaggcggac | attgcagtga | gccaaagatg | caccattgca | ctccagcctg | 127740 |
| ggcaacacag | tgagactcca | tctcaaaaaa | aaaaaaaaagt | cttctatcaa | tgctaatttt | 127800 |
| cctgatcatg | atcattgcat | tgtggatatg | taagagaatg | atcttaggaa | tttagcggta | 127860 |
| aagaggcctc | atgtctgcaa | cttcagggaat | ataaatatgt | aggtagatac | ataaaataca | 127920 |
| tatgcataatg | tataatagat | gtccatatat | gtataagtg | acatgtccac | atagagtgga | 127980 |
| tgtgcataca | cacaagtgca | cctgtatata | tgcaagtcta | tatccataca | tttatatgta | 128040 |
| tgtatgtg | tgtgtgtgtc | tgtgtagaca | gcgagaaaga | ttagcaaat | gtgcacaaat | 128100 |
| gttaataaat | gggaaatcta | gggttaaagg | atataatcat | tatttgtctt | tctctgcaac | 128160 |
| ttttcataag | tttaaaacttc | caaataataa | ataggaggga | actaggagg | tcaaagcatt | 128220 |
| gccaggtgtg | gcagagctgg | gacatgagct | caaggccacc | tccaggttag | tggccttgaa | 128280 |
| gttccccatgc | ccaggacccc | taacctctct | tggggctcca | ccatctggct | ctgctctgac | 128340 |
| tgtgtccagt | gccaccccac | cctggcctct | taoggttgac | taccacgctg | acagggcccg | 128400 |
| agcgtcgtct | gccattggag | ttctgcacct | cacagaagta | gaagccagta | tcagccctag | 128460 |
| tggccaaagt | cagccggagg | gtatgggagt | gggcctcctc | cagcaggaca | tggttcttgt | 128520 |
| accagctgta | gcggagatca | ctgggtgctt | cgttgggtgt | gttgcagact | agtgctcactg | 128580 |
| tctggttctc | caggatggga | ccctgctggc | tcacctggac | ctcagccact | gcaaggcgag | 128640 |
| catagggagt | gctggggggg | cccgctccaa | ttccagcccc | atgccccatag | agtccccggg | 128700 |
| tctcagccaa | tcagcctcaa | tctctgcaca | tcctacacca | cccgctctgt | gtcacagggg | 128760 |
| agcctcctg | agtgctctgc | ctctctttgt | cctgtctaac | aggatccccc | tgcacacccc | 128820 |
| cccttctgtg | gggtattgtc | acctgtctca | tggttaaggca | gggctgggac | tcacacacctg | 128880 |
| catccaggaa | gcactggggt | atgcagttg | ctggccctgc | cctgcccctg | ctccccctcg | 128940 |
| tccctagggc | ctgagctcct | ccttatctc | tggcccaatc | gcaaccttcc | caagaactca | 129000 |
| ctgaagatgt | ggaggctgat | gggggggtg | accaaaagac | ccacgcggtt | ctcagcttgg | 129060 |
| caggtgtaga | gcgcagcctc | gctccaggct | gctcggggca | ggtgcagcac | accagctctt | 129120 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| gtttggaggc | gtaccccatc | cttgagccac | ttaatggaac | tgactgcagg | gtagctgctg | 129180 |
| ttcacctggc | aggtgagtg | gaccagctca | cctggaagga | tgttctctcc | cgaggggctg | 129240 |
| agggagatct | tcacacccct | gggggcatct | gcaagtcaca | gtagggggta | ttgggttaag | 129300 |
| tgcctgggga | gggcagagga | tggcacactt | cttcttgccc | ccctttaaaa | ctgcagtcct | 129360 |
| aaggaagtat | gcccagataa | aacagcagtc | cccttcaatc | ctcaccagg | gcattgctctg | 129420 |
| tctgcccatg | tcctgtccctt | tcctcgccccc | tgcgcctgat | gcatctctat | gcccattgag | 129480 |
| agctgcatcat | gtgacgcttg | gccagcgctc | agcctacccc | accagctgta | gttttctgct | 129540 |
| tcccaatgct | gtccattgct | cctgctctat | ttgggatgag | ctccacacac | agggtcattg | 129600 |
| gtaccccatc | gctagcttta | gtggcctgtg | gaagctattg | gtaagggacc | aactacctag | 129660 |
| tgggaggggg | ccaaaggcag | catcaaaacta | gctctgaaaa | tagttaccct | gtttgtgaag | 129720 |
| aagaggccaa | caacacaaa | aactgcattt | ccttaaatct | tgttccaaag | cctctctctg | 129780 |
| tgtattcaag | tgttttatct | ttattttttt | agaaagagga | tctggctcag | tcaggccaag | 129840 |
| tggagtcgag | aggcacaaat | atagctcact | gcagcctgaa | actcctgggc | tcaagtgatc | 129900 |
| ctctctctct | agcctccccg | ggaactggga | ctacaggtgc | aagccaccac | atccagctaa | 129960 |
| tttttgtttt | tttaattttt | tgtagagaca | gggtctcact | atgttgcccc | ggctggtctc | 130020 |
| gaactccttg | cctcaagcaa | ttctctctgc | ttggcctccc | aaagcactgg | ggttacaggt | 130080 |
| gtgagccaat | gtatccggcc | tcaagtgttt | aatttctgct | agtcactctt | ctaaatcctt | 130140 |
| gacctgggtc | atctccttta | ttgtttttgt | ttgtttttgt | gttcttttgt | ttgagacaca | 130200 |
| gtctcactct | gtcaccagg | ctggagtgca | gtggcacaaat | cacaactcaa | tacaactcca | 130260 |
| cccccggggt | tcaagcaatc | ctagtgcctc | agcctcctta | gtaactggga | ttacaggcat | 130320 |
| atgcccaccac | acccggtctaa | tttttctatt | tttagtagag | acgaggtttc | accatgttgg | 130380 |
| ccaagctggg | cttgaatccc | tggcgctcaag | tgatccaccc | gcctcagcct | cccaaagtga | 130440 |
| tgggattaca | ggcataagcc | accgcgcccg | gcccattctc | tttaattttt | atcataaact | 130500 |
| tgtgagacag | gaaccatcta | tttgttatct | cgtcatagac | tgagaaaaa | gagccaggca | 130560 |
| ggaaagggat | aaatcccacc | tcccagcact | cctccagcac | gggggtaact | ccggggcact | 130620 |
| tgtctccaac | cctgaagctg | catgggagct | cctgttcccc | aaacctttgt | gtctcccttt | 130680 |
| ctctgccccag | cctggcctca | aggacaagcc | ccttcgaagc | agtaggaggt | tgagggaacc | 130740 |
| attttaacgaa | gtccttgggg | ctcagcagtg | tctctccact | tgcccttacc | tggaggccca | 130800 |
| ggagcagcct | gggttttgc | caggagcaag | gggtccgctt | ctgtgggctg | gagaggggct | 130860 |
| ggtttctgtc | aggagcaaca | gatcgctca | gccacaaggg | tgtctcccca | gacaaactca | 130920 |
| acttcacttg | gaggtgaatc | tcgctctgag | ccctgtgatt | ggccacggag | agctggcagc | 130980 |
| gcaggatccg | gccttggtcc | tggccatgtg | tggccatgtg | gaggggtctc | aggtggccga | 131040 |
| cgccggtggg | ctcaaaactc | tggctgttga | aggtgacaga | gcgagcagg | tcttggcctt | 131100 |
| gccactgcag | tctcagctgc | tcctgcaggc | atacgtagg | agtggagcag | tgtgaagtcca | 131160 |
| cctctgtgcc | ctcgagaagc | tccaccgggg | aggcaatgg | gggacacctg | ggctcctctg | 131220 |
| aggacagaga | cagcagtgct | caggaccctc | ttttgccacc | cctgagatcc | ctcgccctgt | 131280 |
| aaaccccagc | tgaggagaga | gccctgggga | gggggctttt | aggggaagga | aagacgcttt | 131340 |
| ctaactccagc | cccacttggg | tgaatacaag | ggagaaccag | gcctgggctt | acgccggggc | 131400 |
| tggaaacagag | gctgagactg | gctgggggta | gattcaggac | aagggctggg | gctgagagcc | 131460 |
| aaggggtcca | gaagcagctt | ggggaatcc | cccggggggc | agccaggcca | ccccacttat | 131520 |
| cacctgttac | tgtgaccaag | gtgcctttca | catctgacca | gcgggtgacc | tcactgatct | 131580 |
| cgaagcgga | gttgtaggaa | ccagagtctc | cgggctgcat | gtccttcagg | agcaggttgc | 131640 |
| acacccctgt | ctcggggttc | ccactgaa | cggtcgggcc | gcgggaagcg | gcctccacca | 131700 |
| gcttggggct | cgcgagatgg | ctcaccacct | gccgctggcc | cagatagctg | tagtacagga | 131760 |
| tggcctgtat | gccctcgggc | acctccacgt | cggcagggaa | gctgaagatg | caggggataa | 131820 |
| gcagccaaaga | cccccttaca | ccctgcacgt | cctggggact | ggagacgccc | catgaggcct | 131880 |
| ggcctggggg | aagaacggca | gggggacaga | ggggagggtg | atacagccct | caggggtgcca | 131940 |
| cagagccctg | cgacctggcc | cagagaaggt | gccccagctg | gcctcccaat | ttctgcctct | 132000 |
| ccccgagaaa | tgcacactta | gagcagccct | tctcagtgcc | ccaggggtca | gtccactgcc | 132060 |
| cgaggtctgt | cagaggtttg | ttaggggtgc | tccaagacag | atgggcagctt | cctgccccag | 132120 |
| tcttggccac | tgcocgattg | ggccctccct | tgaacctcag | aaaccaactc | gtgaccagag | 132180 |
| catagtctaa | tctgccaaag | gctggaactta | tgaaggctgg | accctgggaag | acaggcacta | 132240 |
| ggcccgcaaa | gcttacctgc | tgggaagaat | gaggccagga | ggagaagctt | gggcaagaag | 132300 |
| cccatacgag | gttctctgtg | tgcctctgtt | gcctaagagg | gtggctgcga | ctgtcgctggc | 132360 |
| tgggctcaca | ggggcctcca | ggggaacacct | tgggcaactt | agccccagca | cctgctagaa | 132420 |
| gtccgagcct | gtgtccccc | ctcctctgct | ggccaaccca | ataagagggc | agggctctta | 132480 |
| aagacctctg | agtcagacac | cagcagagag | ccaggaggcc | acgttccccc | ctcaggctgt | 132540 |
| gcccaggaa | gccctcactt | ggtggcctgc | ctcagaaaag | cctgtgtgtc | ccttggccct | 132600 |

| | | | | | | |
|-------------|-------------|------------|------------|-------------|-------------|--------|
| atcccaagtt | ctgctttccc | agccctcaa | ggataccacc | ctaaggcaga | tgaacagctg | 132660 |
| ttctccctc | ttccccactt | ccctgcccc | tccccaccac | ccaaaccaac | aggaactgga | 132720 |
| gccagagat | gccaggttac | tcactccaag | acaccagct | agaatgatgg | ttctctcctg | 132780 |
| aggcttgct | ccctaccctc | gccttactaa | ctatagacca | taatggggct | ttactgaact | 132840 |
| tgccgaagt | ctgcctttaa | cagtcactcc | cctgctcaaa | aacctctgt | ggctcccat | 132900 |
| tgcccgtag | atgtgaaaa | tcctcatttc | ctgccccag | ctctgtcccc | attccccgt | 132960 |
| ctccgcagac | ccactctggg | ggcagttctg | tctgctcaag | ggctccctag | ctgccagct | 133020 |
| ctatctccac | cacagataat | ctttgcctgc | tgaaacctta | ttcaaccttc | aaggagcagc | 133080 |
| atgaattgg | cttccagctg | gaacttctcc | tttgaggttc | ctgtagctac | ccagagcta | 133140 |
| tctctatttt | cottgccttg | ttttttacag | cttgtagag | ccattttctc | aggacctaga | 133200 |
| actgaagata | tgtgccccat | agcagtgagg | agcctaccag | gcattcagca | aaccccttag | 133260 |
| tgactaagag | aggggtgagg | tccttagggg | ttcagagctg | aggttcagag | ttggagtggg | 133320 |
| gaggtggcaa | ggcaagtcta | gggttgaaag | gtagcatgag | agcgctgtgg | aacacatacc | 133380 |
| ccacaaatat | gagctcaatg | tgtgcggagt | gtaccatatc | caaaaaggca | ggccctcaac | 133440 |
| catggagtgc | ccttggtcag | gcagtgtcta | aggggtacca | tagacctgag | cccaaaagga | 133500 |
| agagatgcc | gaaacacata | taagtgaaac | tataaaactc | ttagaagaaa | acaggtgaaa | 133560 |
| atcttcatga | ccttggtatta | ggcaatgctt | tcttaagtat | gaaattaaaa | gcacgaaaaa | 133620 |
| aaaaaatagg | taagttggac | tttcaataaa | tttaaaactt | ggccagacac | agtggctcat | 133680 |
| gcctgtgaac | ccaacacttt | gggagggctg | ggcaggagga | tcactttagc | ccaggagtct | 133740 |
| aagacgagcc | tgggcaatac | tgcaagactc | tgtctctacg | aaaaattaaa | aaacaggcct | 133800 |
| gtggtccag | ctactccag | ggctgaggtg | ggaagatcgc | ttgagcccg | gggaggggct | 133860 |
| gaggctgcag | tgagccatga | ttgcaccact | gcactccagc | ctgggtgaca | ggccaagcag | 133920 |
| ctgtctgaaa | agaccaaaa | acagctgggt | gcggtggctc | acgcctgtaa | tccgcact | 133980 |
| ttggggagct | gaggcatgca | gatcacaaag | tcaagagatt | gagaccatcc | tggccaacat | 134040 |
| gggtgaacccc | cgtctctact | aaaaatacaa | aattagctgg | gtgtggtggt | gtgcacctct | 134100 |
| agtcctagct | actcctagct | ctcaggcagg | agaatcacct | gaaccaggga | ggcgggggtc | 134160 |
| acggtgagcc | aggatcacgc | cactgtactc | cagcctgggt | acagagttag | actcttctgt | 134220 |
| ctcaaaaaaa | aaaaaaaaaa | aaagggatat | gaatcaaaag | acactatcca | gagagtgaat | 134280 |
| aaaaggaggga | caaccacacg | aatggggaga | aatatattgt | atactatttat | ctgataaaag | 134340 |
| actaatatcc | agaatatata | aagaattcct | acaatgaaca | acaacaacca | ccaaaâaaac | 134400 |
| catgaaatcc | aactccaaaa | atgggcaaaa | cacttgaata | aacatttctt | caaagaagat | 134460 |
| atataactgg | ctgataaggga | catgaaacat | tgctcaacat | cacttaggcat | taggaaatgc | 134520 |
| aaatcaaaac | caaacacacg | tgagatacca | cttcacatcc | gttagaatgg | ctattcaca | 134580 |
| acaaacaaag | caacacagaa | aacaataaat | attggtgagg | atgcgaagtt | gaattctctg | 134640 |
| tgtattgtct | tggggaatat | aaaatgggtc | cgtcactgtg | gaaaacaatt | gggtcattcc | 134700 |
| tcaaaaagtc | aacataggat | taccatatga | tccagcaatt | ccaactcctag | gtatatacc | 134760 |
| aaaagtactg | aaaacaggga | ctctaacaga | gtacaccaat | gttcacggca | gcactattcc | 134820 |
| actaaaaggt | ggaacaggt | caagtgtcca | tcagtgaatg | gabtgtggata | aacaaactgt | 134880 |
| ggatgtgaca | tacaatggaa | tatcaatcgg | ccataaggag | gaatgaattc | taacatatgt | 134940 |
| gaaccttgaa | aacattatgc | tcagtgaaac | cagccagaca | caaaagggca | aattattgtag | 135000 |
| ggttccaatt | acatgaaata | tctagactac | gtatatccag | agactgaaag | tagaatagga | 135060 |
| tagaggtaac | caggggctgc | aggggagggg | agctaatgtt | taatgatgtc | tgagtctctc | 135120 |
| agataatgaa | aaagtctctg | aaatagtggt | gatggttaca | caatatgtga | tgtgtacctc | 135180 |
| atgtctatggg | gtgtacactt | aaagacagtt | aaaacagtaa | attttacatt | atgtatatatt | 135240 |
| caccacatac | acacctatgt | tgccaacttt | gcaatcctcc | ctggctcctaa | atgtcctact | 135300 |
| ggccaagtga | acgaggaggc | tggaagtggg | gacaggaaac | tcatgacctc | ccagctccca | 135360 |
| gccactccgc | ctcaggggct | gggctcagca | gattccaatg | actaccaggg | gtcacacctg | 135420 |
| ggaagggggt | gagccgaggc | ccaggggcag | ctaggctgac | cagggtgggac | ctgctcctgt | 135480 |
| gcagaaggca | gaaggtgccc | cagcaggggg | cacagtacag | ggcgggattg | ggaagggaag | 135540 |
| gacacgcctc | cccaggggag | ccagccctct | cgcaggctgc | tggagtggag | tgactctggc | 135600 |
| atttatggag | gcccaaggac | tctctcagga | agccctaggc | agccctaggg | ctcctcctct | 135660 |
| tctgggaaga | tgcaccccca | gcctccacac | cagggtcttg | gccactggag | aatgatatag | 135720 |
| ctggggccct | gggacattgga | cacctcacgc | tgaagaaaag | cagcctgtgt | ggcacactgg | 135780 |
| gggtcagatg | tgtccctggc | cacaggggat | gtcagggtca | gctctgtcat | ggccagggga | 135840 |
| gctactctgt | cccagctccc | ggttctctcc | catttggggt | cctgaaaagg | gcaactgtga | 135900 |
| acctgatgga | agaatgggtg | ggttcttgac | acagcgacc | tggacaaggg | gcggggggag | 135960 |
| gaccctttcc | aggaccactc | ccatcacata | atgtagaggc | cacctatgct | tagcccggtc | 136020 |
| ctaaccctaa | aggggtcagc | cccaccggaa | tccagcctat | tggtcagcc | tgtcaccaca | 136080 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| aaggccagct | tcagcccaga | taactgttct | ggaaacagaa | agagcaggga | ccgctcagaa | 136140 |
| aggagatctc | tgtccctggt | tgaaagcctg | gagttgaggg | gacagtgcgc | cgcgccccgc | 136200 |
| ccccgcgaac | ttaggttgca | gctgtggcct | agttagcacg | cagcgccccc | tgggtggtcga | 136260 |
| gggggaattg | cgggtcccgg | gaaagggggc | ggtgtgccag | caacaggggag | caggcagctc | 136320 |
| tgcagccctg | aaccatccct | cccttgggtg | actcttttgg | aatcatttgt | tccccagaca | 136380 |
| ggaggttccct | gaggttcata | tttgggtctc | caagtcttgg | gtgctctgaa | gacaggattt | 136440 |
| taaatcccca | ctcctactat | tggtatgtgt | tgcatacagg | tggcttgagc | tggcctggta | 136500 |
| cacagtgggc | actcacgttt | gctgcctgtt | tgagaccaag | tgcttcagga | ggtcttggag | 136560 |
| ggttggtctg | ggccccaagt | ccctgacctc | tgattccaga | ggccaagttt | agctggggaa | 136620 |
| gaaagggcag | aggcagtttc | cctatggaca | gctaggcccg | ggtgtaggat | tcagtttctg | 136680 |
| tttcttgaca | ccaaggcttc | tcccacaact | ccccattggg | ctagagaagg | aagaacacag | 136740 |
| ggtgacatg | ccagctggag | gttactggcc | cacagatagg | gagtcagggt | ccggtagggc | 136800 |
| aattcctgga | gcagattatg | gtcaaaatag | tggaaatccc | caatcacagg | ccaaatgttt | 136860 |
| aattctcaga | gccatagaat | ccataactaa | tgcttattgg | ctttgcacatg | ggcagtagaa | 136920 |
| atttccact | tcattcctta | atctggctct | aaatgcttct | ggctggagtg | ctcactctcc | 136980 |
| aaactgtgct | ggacagcacc | agaagccctc | ctgtggacgg | acgaagtggg | gatggatgaa | 137040 |
| gtggaattgt | gctgggggtta | gagtaaggag | agattatcgt | tgggctgggc | tgggctgag | 137100 |
| ttgggtatg | tgatcaggtt | ggattctggg | aagcaaatga | aggctgggat | tggggtggac | 137160 |
| attgaggttg | agtgatgggt | gggggtaagg | gtgaagggtt | gagttgggtg | caggtgatgg | 137220 |
| ttaagatacg | gtggaggctg | ggttgagatg | aggatgatag | agtcggagtt | gtggttaggg | 137280 |
| ttgggtatg | ggatggttgg | gattagatga | gatgagtaaa | tggttagggt | tggggtgcaa | 137340 |
| gtgaggttag | ggtgagataa | ggttggaata | ggggctgggg | ctggggctga | ggttagggtca | 137400 |
| gagcagggtg | agtgctggga | ctgggattgg | gatgcaagtt | ggaaggattt | gggggtgggt | 137460 |
| gaaggggttc | agttgagctc | atctcttgat | agtgtcttgg | ggttgggggt | gtttgggggc | 137520 |
| caccactcgc | accagatctg | agcccccccc | cgaccctctc | ccctatcccc | ctcagccagt | 137580 |
| ttcagccag | cccgcctcct | aatgctccac | tcaccctctg | ggcccaggga | caggggacag | 137640 |
| gggtacctgc | tccacagaag | gaagtggctg | cggcgggtgct | ggacctgcgg | agaggagaa | 137700 |
| aggaaggacg | gccaaagctc | cctgggtgac | ctggctcccc | agggctctgc | cggctcaaga | 137760 |
| gagaaggatc | ccgtatcagg | cggtactctc | tctttcccaa | agcctcagct | ctactgtcca | 137820 |
| accacagagg | tggtcaggga | ggcagctgca | ggccttgctc | aatgccaaaa | cgggaaagac | 137880 |
| ctccataggg | gaaggccctc | ggcaaggcca | gggacttagg | gactccagca | agcagaagtg | 137940 |
| ggaccgctgc | aacgtgtgag | ctcccccagg | caaagtgaga | aatggagtgg | ggagactcca | 138000 |
| ttcaccagcc | aaatccacac | cccactctct | ctgagccctc | tagggaggct | gggggagggt | 138060 |
| gaagggggct | tctgtccacac | agctctctc | cctgacacct | gagagggagc | cgcgccagc | 138120 |
| ctgggggtgg | gatgcactac | acgatgccca | gaccaaggca | gttattctcc | aagccattca | 138180 |
| ggagcctccc | tgtaccactg | agtactctta | agaaccccca | agaggcagtc | ccgtgttgtg | 138240 |
| ggggacataa | ggccccaatg | tccaggacac | tctcagggtt | ttctcttttc | cctctcactg | 138300 |
| tctgtacgt | tttttggttt | ttggctctct | gttgttggtt | ttttcttttt | tgttttgttt | 138360 |
| ttgcttttga | gatggagtct | cactctgtcg | tccaggctgg | agtgccgttg | cacaactctca | 138420 |
| gctcacggca | acctccacct | cccagttcca | agtgtatcat | ttgctctcag | ctctcgagta | 138480 |
| gctgggtata | caggcatgca | ccaccacatc | cggctaattt | ttgtattttt | ggtagagacg | 138540 |
| gggtttcacc | atggtgtcca | gatttggtct | gaactcctga | ccccaaagta | tctgctctgc | 138600 |
| tcggcctccc | aaagttctgg | gattcacagag | gtgagccgct | gcaccggccc | actatcctat | 138660 |
| accttcaccc | ccactctggg | agcgaaagg | aaagcccca | ccacagctag | ttactgttca | 138720 |
| attactgcag | cccacttat | tgagggccta | gtttgtgcag | gcacttgaca | gttgaagcag | 138780 |
| atgttaattg | ctccaagcaa | tcccagatag | aagcaacaag | actttttttg | gtttttgttt | 138840 |
| gtttgtttgt | ttttgttttt | tgagacagtg | tcttgcctct | tcgcccagtc | tgggggtgct | 138900 |
| tgtgtcaact | ttggcttact | tgtgctcact | gcagcttcga | aattctaggc | ttaaagcaatc | 138960 |
| ctcctgtctc | agcctcctaa | gaagccggga | ttacaggcgc | ttgccactat | gcaaggctaa | 139020 |
| tttttttaag | tatttttgat | agacggagtc | tccctatgtt | ccccaggctg | tctcaaacac | 139080 |
| ctggggtcaa | gtgactctcc | cacctcgacc | tcccaaagtg | ctggaaactag | agggcgtaacg | 139140 |
| caccactcct | ggcctatttt | ttttttaatt | tcacaactca | aacttaattc | aactcagttc | 139200 |
| ctgcctacct | gtactgggtg | ggagctggcc | acagcaaatt | agggccctgt | accttcaggg | 139260 |
| caaggccacg | tgtagcaggt | taaaaggatg | tgaaccccta | aatcagggtg | gatccagaat | 139320 |
| cgcaggccat | gggtgcccaa | agcagatgtc | tggtgacatt | ccacctgtaa | tgcctcagge | 139380 |
| taacagata | ttaggtctct | atcactctgt | tctcttttat | agctcctgtg | tcactcctca | 139440 |
| tgcttggggc | atttctctct | tcggccaaaa | cacaaaagggt | tcactccatt | acttctctcc | 139500 |
| tcaacagctg | gtccggagac | accagctctc | aggcctgtgg | ggttgtgaca | catgggtacc | 139560 |

| | | | | | | |
|-------------|-------------|-------------|--------------|-------------|-------------|--------|
| aatccttcag | tccactggga | ctctatacat | ccacccttg | gcttcatggt | gggaaagcat | 139620 |
| cccttgactt | tgaccttagt | catatgactt | gctctggcca | atggatgtgc | acggacatga | 139680 |
| cacaaactaa | gtcttgaaac | gtacttaagc | agtttgcctt | tccccctgaa | ttctctgtcat | 139740 |
| tgccatgcga | ggggcaggct | ccaactaaac | tgctgggtcca | aagaggatga | cgaacacatg | 139800 |
| cagatgacct | gaatcagacc | catggattga | aaacaaactc | agctgagccc | agcctacgtc | 139860 |
| caccagacca | gtcaacctgt | ggatgcgtga | atttatttgt | ggatgctgct | gagaattttg | 139920 |
| tggctacatt | agcaagatga | tacaaggcct | aagtcaccaga | acaacacacc | cagaacttgc | 139980 |
| ttacctttcc | ttagcatgag | gagagcaaac | acttgtctac | cttgattagt | caggggagcac | 140040 |
| tgcttccgtg | catcttcctg | agtatcacgc | aaactaggta | aataaataaa | aataaactagg | 140100 |
| taggtctggg | acggtggctc | acgcctgtaa | tctcagcact | ttaggaggcc | gaggtggggc | 140160 |
| gatcgcttga | ggccaggagt | tcaagaccag | cctgggtcaac | gtggcgaaac | cctgtctcta | 140220 |
| cgaaaaatac | aaaaattagc | tgggcctggt | ggcaggcgcc | tgtagtctca | gctactcagg | 140280 |
| aggtcgaggc | acgagaatcg | attgaaccog | ggaggtggag | attgcagtga | gccgagatca | 140340 |
| cactactgca | cctcagcctg | gatgacagag | cgagactctg | cctcaaaaata | tttttaaaaaa | 140400 |
| tgtaatttct | cattaggtca | cattcatcta | tgtagtctca | ataacaatta | ttctttttct | 140460 |
| tttttttttt | tttttttttg | agatagggtc | tcactgtcac | ccaggctgga | caggctggag | 140520 |
| tgcaatggca | caatctcggc | tcactgtcaac | tttgacctcc | taggctcaag | tgatcctcct | 140580 |
| gctctcagct | cccagattgc | tgggactaca | ggtagtagtacc | accagacgca | gccaattttt | 140640 |
| gtattttttt | gtagagatgg | ggcttcacca | tgttgcccag | gctggtctca | cactcctcag | 140700 |
| agttcccgat | caagtgatcc | accacacctg | gcctcccaaa | gtgctgggat | tacaggtagt | 140760 |
| agccaccata | cccagtgaa | tattctcggt | ccagatagga | aaaccaagtc | ataggggaggt | 140820 |
| tagagaattt | gccaaagaca | aaactttttg | gttgaaaaaa | aataagtttt | gctacaagta | 140880 |
| tagaaaacac | caataaacgg | tggttttaaa | aaaatagaag | gttttaccoc | tctccctcaa | 140940 |
| gtaagtgttg | gcataccaaga | tgatatgaca | actccacaat | catgaaacta | gatccctttt | 141000 |
| tttttttttg | ctcagtcatt | gtcaatgggc | tgcttccagt | cattgtccaa | agtggtctgt | 141060 |
| tgctgtccag | ccactgcgat | tgcatctcgg | aaagcaggat | ggaggaaagg | acaaagttagg | 141120 |
| ccatgcccctc | tactttaaga | taactcctta | aattgctgat | gtggtgggtc | actcctgcaa | 141180 |
| taccagccac | tcaggaggct | gaagcaggag | gttcacttga | accaggagac | ttgaggctgc | 141240 |
| agtgaactat | aattgtgtca | gccttgatcca | gcctgagtga | cagagtgaga | tcttgtctct | 141300 |
| taacaacaaac | aaaaaagcta | attcttttaa | ttgtacaact | tattgtgctg | tatatcccaa | 141360 |
| tagaagctag | aagttagtca | catggtctta | agtcctttatt | ctaaacaggtg | atgtgcctcga | 141420 |
| gtgaatcctg | ggggcccagt | gctaagaagt | agagggagat | gaagcctctg | cctcctgttg | 141480 |
| agcaaccagg | accttaattc | agtcattcat | tccttggacc | taccagatt | ccagcctcag | 141540 |
| gccctcgccc | accctcctgc | cagtatctct | cgagagcctc | ctggcctccc | cagagggtgc | 141600 |
| cagagccctc | tgcttctctc | gccttcagac | ccacttgcgt | tgtctcaggt | ggggacaaact | 141660 |
| cagtcctatag | aggcccaagg | gagatctttg | agaccatac | tgtagtcagg | ttggggcaga | 141720 |
| tgggagtcac | gtagctcact | ggctaaaaag | ctaggctcgt | catagagctg | gggtcaaatc | 141780 |
| caggcatatc | catcacctgc | tctacgatcc | caggagtcac | tcaatgactc | agaagctaga | 141840 |
| gtcctcttga | aagcttgtat | gaagagttcc | cagggccttg | aaacataaaa | gagctcttag | 141900 |
| aatgtgtgct | ctgctgttgc | ttctcttatt | gggtatctga | aggccgtact | ctcctctctc | 141960 |
| ctccaccoga | ggtgacctct | actctcatag | ccacacccctg | gacctctcat | catctcagcta | 142020 |
| cctcaccagc | tctgtcatct | gtccctcttg | ctctcctctc | ctgttcatgc | atcttgtact | 142080 |
| caccacattc | tgctccttag | tgtttcaagg | ctctgggggt | ttcagagaca | ctgaagggac | 142140 |
| ctccctctct | aacacaaaca | caaggtctag | gtggaaatgac | ccaactaagg | accggctctg | 142200 |
| aagccagaga | ctccaggga | aagtcacaaa | gcccaaggat | gcccggtaca | agaaagttaa | 142260 |
| aagaccatg | tacatgtcct | cccgttttat | tccttgctca | gggtctgggc | atacagtgga | 142320 |
| acacatcgag | ttcccaaaag | acacogctg | tacagagtca | gatggagtta | agaacattta | 142380 |
| gcgggtctgg | cgcggtggct | cacgctgtga | gtcccagcta | ctcgggaggc | tgagggcagga | 142440 |
| gaatggcggtg | aaccogggag | gcggagcttg | cagtgagccg | agatcgccgc | gctgcactcc | 142500 |
| agcctggggc | acagagacag | actctgtctc | aaaaaaaaaa | aaaaagaata | tttagtctgg | 142560 |
| gcacgggaag | ttatgcttgt | aatccagca | ctttggggagg | ctgaggtggg | cggataacct | 142620 |
| gaggtcgaaa | gtttgagacc | agcctggcca | acatggtgaa | accccatctc | tactaaaaat | 142680 |
| acaaaaatta | gccagggac | tgtaatccca | gctaactcagg | aggctgagcc | aggaataatca | 142740 |
| cttgaaccca | ggaggtggcg | gtcacagtga | gccaaagattg | tgccagtgca | ctccagcctg | 142800 |
| ggtgacagag | caagactcca | tctaaacaca | cacacgcaca | cgcacacata | tttaaaccca | 142860 |
| cacaccaaca | tctaggttcaa | gtaggtggac | tctctgccat | tctctgccat | tctctggccc | 142920 |
| tccaatacca | ctgagagcac | agtaagcaaa | gggaaaagga | gacagaagg | ctgggaacag | 142980 |
| gatggctggg | ggatggggaag | tatccactgc | acaggatttt | gatttaattc | tagaagatag | 143040 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| aaagaggagg | gatcacggttc | aggaacagat | gcgggttaaag | gaaaccagag | ccaaagcacg | 143100 |
| ctgagagaaa | gctgccccag | aggccggagc | agaagtggac | tctctacagg | gactcaatac | 143160 |
| accocaaagg | gttggttagct | ggcacacgta | cctctccacc | ccacactgta | acactgcagg | 143220 |
| gcagaggaaa | taccctagggt | gagttccaggt | atcaatcaac | ctgcccctttg | ttcataaaga | 143280 |
| tgaatttggtt | accaggtattt | accagacagg | tgagggaagac | caacacaaag | agaaagatcc | 143340 |
| agaaacaac | agccacaggcg | cattgggtc | cgctgtaat | cccagcacte | tgggaggcca | 143400 |
| agatgggtgg | atcacctgag | gtcaggagtt | caagaccagc | cttgccaaca | tggtgaaacc | 143460 |
| ctgtctctac | taaaaataca | aaaattagct | gggtgtgggtg | gggttctcct | gcctcagcct | 143520 |
| cccagtagc | tgaggaggctg | atagaggagg | ttcacttgaa | cccaggagct | cgaggctgca | 143580 |
| ttgaactatg | attatgtcac | tacattccag | cctgagtgac | agagtgagat | ctgtctctct | 143640 |
| aacaacaaca | aaaaagggtaa | ttctttaaat | tgacacactac | attgtctgctt | atatccact | 143700 |
| ggctagagg | tagtcacatg | gtctcatgtc | tttattctaa | caggtgatgt | gcccgagtta | 143760 |
| atcctggggc | cccagtgcta | agaagtagg | ggagatgaag | cctctgccc | ctgcttgaac | 143820 |
| cctagaggta | gaggttgca | agagcggaga | tcactgccact | gcactccagc | ctgggtgaca | 143880 |
| gagtgagact | ccatctcaaa | aaataaaaaa | taaataaata | aaataaata | taataaata | 143940 |
| aataaaaaac | tacagcaaa | aaacaaagac | tattgcaccg | ttactgtggg | cctggcagta | 144000 |
| ttatgacaac | tttttttttt | tttttttttt | ttggagatgg | agttattttt | tgtaaccag | 144060 |
| gctggagtg | agtgagtgca | ttctcagctca | ctgcaacctc | tgccctctgg | gttcaagcga | 144120 |
| ttctcctgct | tcagactccc | gagtagctgg | tattataggc | acatcccacc | acacccggct | 144180 |
| aatttttgta | tttttagtag | agaccgggtt | acaacatggt | ttccactgt | tgccagctct | 144240 |
| agtatcaaac | tcctgacctc | aggtgatcca | ccgcctcag | cctcccaag | tgccaggatt | 144300 |
| acagggcatga | gccaccacac | ccatccagtg | ttctgacaat | tttatactta | tttaactcatt | 144360 |
| tagcaaccct | ctgagctaga | gtctattgtt | atctccactt | acaggaactt | gagtccaga | 144420 |
| atggtacagt | aaaataacct | tgaccgaggt | ttacatggct | ggtaagtggc | agagaaatga | 144480 |
| aacttgaacc | caggcaacct | gtttctctgag | ttgagctctg | aactactcag | ccatactgcc | 144540 |
| tattttatta | tttatttttt | tactacttta | cttatttatt | gagatggggg | cagccggcga | 144600 |
| tggtgggtta | cgctctgta | cccaccactt | tgggagggctg | aggggaaacg | atcacttgag | 144660 |
| gccagcagtt | cgagaccag | ctggacaaca | aggagacagg | gtcttgtctc | gtcaccagg | 144720 |
| ctggagtgca | atggtgcaat | catggtctac | tgacgcttcg | acctcccagg | ctcaagcgat | 144780 |
| cctcccacct | cagctcccca | agtagctggg | accataggca | cccactatca | tgctccagta | 144840 |
| atcgaaaaaa | aaaaattata | tagagatggg | ggtctcacta | tattgctatg | actggtctct | 144900 |
| aactcctggg | cttaagcaat | ccaccacctc | tgccctccca | aagtactaga | attacagggtg | 144960 |
| tgagccaccg | tgcttgacca | atactctctc | tttaataaac | taaaacaaaa | gtgaacaaac | 145020 |
| caatcccggg | gggaacagat | aattcaagg | atacaagaa | acttatgaaa | agaaagattc | 145080 |
| tagtgccaat | atgttcataa | aaacaaaaaa | agcagctatg | aaaagaaagc | aaaaaataat | 145140 |
| tcttggaat | ttaaaaataga | attgtctgaa | taaaaagaaa | ttcaatagaa | aaggggtgtga | 145200 |
| acatctgtgg | ttttgggttc | caaaattagt | tcactctctc | ttgggtaaaa | ataccctgat | 145260 |
| tttcctgctg | tattctcaag | ccctgtgggt | tggtggaatt | gacacctctt | ccatttactg | 145320 |
| cacgagagaa | gcattgtgct | gctcacagtt | cattatgagg | agccagcagg | taaagacaac | 145380 |
| actgagggca | gatgaaagat | atatactgac | cctgggtcct | tggtggacatc | attgagctcc | 145440 |
| tggtatcaagc | cttacctgaa | gctaaaggga | tgagtgtctc | tcactgttaa | tgccactgta | 145500 |
| tcaaattcca | tagaaggact | ccaattatcc | ttgcctcagt | tggtgttcca | gcctcggag | 145560 |
| gagctaccat | caagggaac | tgccagacct | tggtgacaag | cccaccacct | gaaaaatggg | 145620 |
| aaaaaaccca | agtcacatca | ttgaactgaa | agagctgcag | ttcccgcgaa | gacaggagag | 145680 |
| aagggaatgct | ggataaacac | aacactact | tcactcact | gcacttgcga | agactccagg | 145740 |
| cacaggaagg | tgggcaccagc | aagtagtctt | ctctggttgac | ataacaaaga | caacatctgt | 145800 |
| ttcttctctc | agaaagaagt | aagaaaggca | attggagaac | cagaaacgac | aaaaatctat | 145860 |
| ccaggaccaa | atagaaagaa | aagcatgatt | tggaaccaat | tggaaccaat | agtggggtgt | 145920 |
| aagaaaaag | aatccatttg | accttgacat | ctgaaacagc | aaatatctct | catcaaggca | 145980 |
| cttaggggta | gaggaaatga | tactattgtt | ccttctcaaa | ggagctatct | ggttactggtg | 146040 |
| ttctgcagta | aatgacgttt | atacaatcat | aacatccgtt | gcttctcaat | tttcagaaac | 146100 |
| aaactgaaga | taagagcatg | aagattttat | tataattacc | gaacagcatg | caaattgtac | 146160 |
| aaactttgac | taaaaaaatg | ccaggcgca | ccaggcgca | gtgggtcacg | ctgtaatcc | 146220 |
| tagcactttc | ggaggccgag | gcgggtggat | cacaagggtca | ggagtttgag | accagcttgg | 146280 |
| ccaactatgt | gaaaccccat | ctctactaaa | aatacaaaaca | tagctaggc | atggtggcac | 146340 |
| gcacctgtag | tcagagctac | atgggaggct | gaggcaggag | aatcacttga | accaggagg | 146400 |
| cggaagttgt | ggtgagccaa | gatcacgcgc | cggaactaca | gcctggtcaa | cagagcaaga | 146460 |
| ctccatctca | aaaaaaagaa | aagaaaagaa | aagaaaagta | gacagcagaa | attagaggga | 146520 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| gatggcaggt | gacctggggc | tgggggaggc | acaatttttc | ttacatatgtg | atagggctca | 146580 |
| agaaattctt | tctaaaattg | atgtatgagg | aacgaaat | taaatataga | ttgttttagaa | 146640 |
| ctataagtag | caccaacaca | ggactccagat | aatcacacag | caaacaaaaac | tgagaaaaa | 146700 |
| gacatggg | gggagatggg | aatgggtgat | ttccattccc | tgcttttaagt | atgaggtgtc | 146760 |
| agtagatgct | ctctctgagtt | actaaattga | gaacacaaaga | taaaagtatat | tgttttagagg | 146820 |
| agtgatgttc | aaactccaga | aaaggctggg | cgcggtgggt | catacctgta | atcctggcac | 146880 |
| tttggggagc | cgaggtgggc | agatcactta | agcccaggag | tttgagacca | gcttaggcac | 146940 |
| catgatgaag | ccctatctct | actaaaaaaa | aaaaaaat | aaaaaatatg | caagcagtcg | 147000 |
| tggtgacac | ctgtagtcct | agtttctcag | gatctcacta | ctgtactcca | gctcggttaa | 147060 |
| tgagagtgag | acctctctct | aaaaaaaaaa | aaaaaaaaaa | aaaagaagaa | gaagaagaaa | 147120 |
| actccagaat | atataaaaaa | gaaaagagg | cggaagagtg | gaactgcggg | gtagaaaaat | 147180 |
| ttgcattttg | tttcatctct | ctgtatatgt | tgaacttttt | ttttaacct | catgtactac | 147240 |
| gtcatatcca | cagcccccaa | ccatccacca | ggcatcaact | gctgtgttta | tatggagggt | 147300 |
| tgagcaatca | ttctgtcctc | cttccagttc | gtcttctctg | actgcagagg | ctagaaaaat | 147360 |
| aaattttatat | caccagattt | ccctccagc | taccttaagt | ctaccacctg | atccagccat | 147420 |
| cattgatctc | ccacttcacc | aaactcggta | ataattgtct | cttctaataa | tattagaatt | 147480 |
| gggactaaga | gacactacac | ctgacccctt | aacaacatgg | gtttgaactg | tgctagtctg | 147540 |
| cttatacaca | gatttttaaa | aaaaataaca | tattgaaaaa | aattttggag | agatgtgaca | 147600 |
| atttgaaaaa | actcacaac | cacatagcta | gaaatatcaa | aaaaaaaaaa | actgaaacag | 147660 |
| agttaggtag | gtcaggaatg | cataaatata | tatgttaatt | ggctgtttat | atttattagta | 147720 |
| gggcttccag | tcaacagaag | gctattagta | attaagattt | tgagtcaaaa | gttatacaca | 147780 |
| gatttttgat | gatgaggagg | attagtgccc | ccaaccccc | tgtttgtcaa | gggttagcta | 147840 |
| tatttgaaat | ctgctggtag | ctggaattgg | gcaactaaat | tcacgagatg | tggaacggtc | 147900 |
| gtatacatgg | agaataaata | aataaataaa | taaataaata | agttatttgt | ggcctggtgc | 147960 |
| agtggtcat | acctgtaaat | cccaacactt | gggaggccaa | ggcaggcgag | tgacttgagg | 148020 |
| tcaggagtg | gagaccagcc | tggccaactt | ggtgaaatac | cgctctctact | aaaaatacaa | 148080 |
| aaatcagctg | ggcgtggtgg | tgtgcacctg | taatcccagc | tgcttggggag | gctgagggat | 148140 |
| gagaatttgt | tgaacctggg | agtcagaggt | tgcagtgagc | cgagattgca | ccactgcact | 148200 |
| ccagcctcgt | cgaaaagagt | ggactctgtc | tcaacaacaa | caacaaaaaa | gcaattgtta | 148260 |
| aaagaaataag | aaaacaagcc | acagatctgg | agaatatatt | tgcaaaaacaa | atatctgatg | 148320 |
| acccaaaaata | cacaaagaa | ctttaaacc | caacactaag | aaatcaacaa | accocattaa | 148380 |
| aaaatggaca | aacggggccag | gcatgggtgg | ggtgggtcac | atctgtaaat | ccagcacttt | 148440 |
| gggagggcca | gcaggggcaga | tcaccttagg | tcaggagttc | tcaactgagc | tgggccaacat | 148500 |
| agtgaaaccg | tctctactga | aaatacaaaa | attagccgac | cgtgtgtgca | gtctctgtg | 148560 |
| gtcccgacta | cttgggtgtc | tgcggcagga | aaatagcttg | aaccaggggag | gttgcaatga | 148620 |
| gctaagatcg | cgccactgca | ctccagcctg | ggcaacacag | tgagactccg | tctcaaaaaa | 148680 |
| aaaaaatggg | caaaagatgt | gaacagacac | ctcatcaaa | aagatataca | gatgcgaagt | 148740 |
| aagcatatga | aaatatgctt | aacctgatgt | cattaaggaa | ttacaaattg | aagcaacacg | 148800 |
| gtaccactaa | acccttatta | aaatgggtcaa | aatccagaat | gctaaaaaca | ccaatgtctg | 148860 |
| gtgaggatgt | ggagtgcacg | gactctcatt | cgttgctggt | ggaaatgtaa | aatggtattg | 148920 |
| ccatttttaa | gacagtttgg | cagtttctta | caaaactaaa | catagtctta | ccatacaact | 148980 |
| taacaatttac | actcttagat | agttacccaa | ttgaggtgaa | aactttatct | cagggccgag | 149040 |
| cacagtggtt | cacacctgta | atcccatcac | tttggggagg | caaagaggaa | ggattgtctg | 149100 |
| agggccaggag | ttctttcttt | ttattttatt | atttattttt | ttcttttga | gacagagttt | 149160 |
| cgctctgtca | cccaggctgg | agtgacgtg | agtgatctca | tctcactgca | atctctgctt | 149220 |
| cccggttcca | agcgattctc | ctgtctcggc | ctctgagtag | ctgctgggat | tacaggtgca | 149280 |
| cgccaccatg | cccagctaat | ttttattatt | ttagtagaga | cagggtttca | ccatgtcagc | 149340 |
| tagactggtc | ttgaactcct | gacctcaggt | gacctgcctg | ctccggtctc | ccaaagtgtc | 149400 |
| gggattacag | gtgtgagcta | ctgtgctcag | ccaaggccag | gagttccaga | ccagtctctg | 149460 |
| caacacagtg | atactctgtc | tctacaaaaa | aaaatttttt | taattagtca | catgtgggtg | 149520 |
| cacacacttg | tagtgtcact | tactggggag | gctgaggctg | aggatcactt | gagcccagg | 149580 |
| agtttgaggt | tgcagtgagc | cacgatatgt | cctctgcact | cagcgtgggt | caacagagta | 149640 |
| agagaaagaa | agagagagag | agagagagag | agaggaagga | aggaagggag | gagaaagaa | 149700 |
| acgaaaagaa | aagaaggaag | gaaggggaag | aagggaggga | aggaagggag | gaaggaagga | 149760 |
| agaaagaaa | aaaggaagga | aaataaaaaa | gaaaagaaaa | gagagagaga | acccctatat | 149820 |
| ccatacaaaa | ccctgcgcga | agctgtttat | agctgtttta | ttcataattg | caaaaaactg | 149880 |
| gaagcaacta | agaagctctc | caatgggtga | aaggataaac | aaactgtggt | atatctgtgc | 149940 |
| aatggagtat | cattaagtaa | taaaaagaa | acttagcaaa | ccacaaaaag | taatatgcat | 150000 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| actaaaatg | catttagtaa | tattaatatg | cattaagcaa | taaaaaaaaga | cgtgtcaagc | 150060 |
| cacaaaaagt | aatatccata | ttactaaatg | aaagaagcca | gtctgaaaag | gctacgtgct | 150120 |
| gtgtgatttc | aactatttga | ctttctggaa | aaggcaaaac | tacagacagt | aaaaagatca | 150180 |
| agggttgccg | ggggttctgg | ggacaggaaa | ggatgaatag | gtggagcagc | ggaatttttag | 150240 |
| gtcagtgata | ttattctata | tgatactata | atgggtggatc | catgatatgc | ttttgtcaaa | 150300 |
| accatagaaa | agtacaacac | aaagagtgtat | tcttaagtga | aactatgtgc | tttagtcaac | 150360 |
| aatgtattga | tattagtcca | ttaatttttaa | caatgtacca | caactaatca | agatgttaac | 150420 |
| aatggggaaa | gtgcgtgtga | ggggggggat | acatgggaac | tctgtactac | agactcaata | 150480 |
| ttttataaat | ctaaaaacttt | tttttttaaat | aaagtctctt | agaaaaaat | agaaataaaa | 150540 |
| taaaaagtca | attttttttt | ttttttttga | gagggaaatct | tgctctgtca | cccaggtctg | 150600 |
| agtgagtg | cacaatctct | gctcactgca | acctctgcct | ccttagttca | agcgattctc | 150660 |
| ctgcctctac | ctcccaggtt | caatcaattc | tcccacctca | gcctcccaag | tagctgggac | 150720 |
| tacaggcatg | cgccaccacg | cctggctaat | gtttgtattt | ttagttagaga | tgggggttcg | 150780 |
| ccatgttgcc | caggctggct | aagaaactct | gacctcaggt | gatccgccc | cctcagcctt | 150840 |
| tcaagagctc | gggatgtgac | gggtgaacca | catgcctggt | ctaaagtgtc | cattttttaa | 150900 |
| aggcagtcctg | gctggggcag | tggctcacgc | ctgtaatccc | agcaccatgg | gagaaggctg | 150960 |
| agcaggggtg | atctctctgag | cccaggagtt | cgaggctgta | gtgtgctatg | atgggtgccac | 151020 |
| tgcactccag | ctctgggtgac | agagtgtgagc | tctgtctcta | aaataaataa | ataaaaatac | 151080 |
| aataaaaattt | taaaaggcaa | tctgcagcaa | gagaagatga | agttagcagg | agagcaggac | 151140 |
| agaggagagt | ccatgaacaa | cagggaagag | aaagtggcca | tgggtggcagc | tcccaggctc | 151200 |
| ctggcatatg | ttcaattccc | tgttcccagc | cctcaggaaag | cccatagctc | cccactctgc | 151260 |
| cctacatata | aacctggat | ccttgacatc | aaactctcct | tccctctatg | tgctctaagt | 151320 |
| agtttttgct | acttgtggaa | catttaacta | acatcattaa | ccaagtggac | ctgtctcaga | 151380 |
| gtggtgttat | gagaagacag | cagccaaaag | acagctgcag | ccaagcacag | tggctcatgc | 151440 |
| gtctaatcct | agcattttgg | gaggccgagg | tgggtggatc | acctgaggtc | aggagttcga | 151500 |
| gaccagatgc | gcacaactgg | cgaaaacccc | tctccactaa | aaatatcaaa | attatccogg | 151560 |
| tgtggtggcg | agcgcctata | atcccagcta | cttgggaggt | tgaggcagga | gaattgcctg | 151620 |
| agccccagggg | gcagaggcgg | cagtgagccg | ggatcgtgcc | acttcactcc | agcctggggtg | 151680 |
| aaagagcaaa | actctgtctc | aaaaaaaaga | aaaaaaaaga | cagctgcaac | aaatgtcaag | 151740 |
| ttctgtgtgt | tttcttttct | tttctttttt | ttctatttaa | ttaatttatt | ttagagtcag | 151800 |
| agcctcccta | tgtcacccag | gtcggagtgc | agtggcacag | tcacagctca | ctagtccctc | 151860 |
| aaactctctg | gctcaggcga | tccttccacc | tccagctcct | tcttagctgg | gactacaggt | 151920 |
| gtgtgccacg | acactctggtc | tgtgtgtttt | cttttctttt | tttttgagac | ggagtctctc | 151980 |
| tctgccaccc | aggctggagc | gactctggct | gatcttggct | cactgcaacc | tctgctcctc | 152040 |
| gggctcaagc | aattctctctg | cctccgcctc | ctgagtagct | gggaatacag | gcgcacacca | 152100 |
| ccatgcccg | ctaatttttg | tatttttagt | agagacgggg | tttcacccat | ttggccagga | 152160 |
| tggctctgat | ctcctgaact | tgtttccac | ctgactcggc | atcccaaagt | gctgggatta | 152220 |
| caggcgtgag | ccactgcacc | ctgcctggct | tgcactgttt | ctgacatact | gtcaaaaagg | 152280 |
| tactcatact | aaatggcaac | acattctcaa | gccccttctc | ttttctctcc | tatctgcttt | 152340 |
| accacacacc | atagctgact | ttaccgtttt | ctccttaaaa | actcctcccc | accttccccc | 152400 |
| aacctatcta | tccacttctt | gtccctccct | caaaaccact | ccattttgat | tctgcatcat | 152460 |
| aaaaaccaag | cagagtttctg | gggagccaga | agcctcctct | tatgatattg | gaagggggtg | 152520 |
| gatgaaattg | agcacacaaa | gccagaatcc | tcttttgtgg | aaagatgggg | gcagtgggga | 152580 |
| gagggaaagca | gggtcatctc | tttccctctc | tcccttccc | tctcttttca | caccacacgc | 152640 |
| tccgcttggg | tgaagctcatc | cttcaatttc | cactccagcc | agatcgactg | agatcgactg | 152700 |
| ccaaatttct | accgcgaatc | ccaaactcca | ctctgagctc | cagaccatc | tactaattgc | 152760 |
| ctagtcaaca | ttttctctgt | gggatcgaat | cagataggat | tatatctctg | tgtgactaac | 152820 |
| agagactgaa | aattcagagt | cctaacaaca | agactttctt | tttctgtgtc | tgaaaagctc | 152880 |
| ggaagtacac | agccctaggc | tgggaatagg | atactgctgg | ggggctccgc | ctccttgcat | 152940 |
| ggtgcccctc | tgtatctccc | aggggtgtgc | cctcaaccac | actgtccaaa | gtgggtgactg | 153000 |
| agggacagct | cctcccactg | acattccagg | caacagaaatg | gagagaagca | ggctgaggaa | 153060 |
| ggaggaaacca | caggcgccct | ccagctgttt | caagggaaggt | tctgggaaga | tgtcaataaa | 153120 |
| caattccggt | tatatctcat | tagccagctc | tagtcacgtg | ccaacaccta | gaagcaagaa | 153180 |
| aggctgggaa | atatactgtt | tctgctgggt | taccgtatgc | atgactaaaa | agagagcgct | 153240 |
| ctgttcttga | gaaaatgaga | atagactggg | tggggggagg | ataaccagca | tggcttaacc | 153300 |
| tgcagcccca | caacctctca | atgcttgatg | tggctgaaat | aaaaagagca | tctccccta | 153360 |
| acatttctct | tcccttagtc | ttgttccctt | tctgtacccc | ataccctagc | agaggggcaa | 153420 |
| ttctgtgggg | tgggctcact | ggcaagggga | ggtctcacc | aaagctcagc | aaaggtgagc | 153480 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| agagagaatt | gtacaaggaa | gctgaatttc | cagctggtct | ggaggaaaca | ctgaaaaccc | 153540 |
| caaacagaag | atgtccaggt | ggagcaaatg | cagctgagag | ttcttgtcca | aggggacaga | 153600 |
| atgtcagcag | agtaggacac | aagggcaggt | ctctactgaa | agccaccagg | gcaaaagtcc | 153660 |
| ccagcccttc | aaccggcttg | ccaccaagc | catcggtcc | actctgcttc | ctgccgtgc | 153720 |
| caccctttgc | acttttgtct | gtaccatctc | ctccaccagg | agtgccttc | ccgctcctcc | 153780 |
| atgttggtata | ctctcagcca | ccccctcagg | tcagatccag | agccaccctcc | tccaagaagt | 153840 |
| ctccacctgt | ctgtctggaa | ccctcacatc | acactgggag | acgaagccaa | aagtgggggg | 153900 |
| gtcagtcacag | aggtgtggga | ggtaggtcca | atcctggctg | gacctggctt | tgggtgccat | 153960 |
| ttgggacagg | gtattgggga | ggttcagcag | gaaggtggga | cctcaaggct | ttctttttat | 154020 |
| gaaggaggag | gaaccagctg | ttgacagagc | agcaacagca | gggagccacc | aaggctgagc | 154080 |
| cttgaacctt | gggccaccag | gctcctgcag | gcaactggga | cgagctcagc | ctctcacagt | 154140 |
| ccattttcca | gggaggaaga | tgatgcagcc | tgaggcaatt | ggccacctgg | aggaagtcca | 154200 |
| gctgggtggg | gcaggcagag | gaattccttg | aaggaaagct | acccctcagc | actggggtag | 154260 |
| agggacacta | gagacagggg | tggcccaggga | gctgccacc | tcctatcccc | atgaataagc | 154320 |
| aggtggggcc | gggtgaccag | ctgggagcca | ccctgggtg | ctcaggtacc | atgcttcttg | 154380 |
| ccagcctccc | tgagctgtgg | cacaggcaag | agcaccggcc | cacggggtct | gcatggggca | 154440 |
| gtgacctgtt | ctctctaagg | ttcagcttct | tcactgttag | aaagagggtg | aaagagctct | 154500 |
| ccccaccatg | gttgtgttgc | gggttaactga | agtgtcaggg | taggggaccc | agtgggggac | 154560 |
| cagcacacct | gaggttcagt | caacaaggcc | cactggtgaa | gaaaacctcg | ccaccctccc | 154620 |
| gtccctctct | gagggcaggtg | gtggaggcac | atgccctctc | catttgttgt | gcatgagtgt | 154680 |
| cgtgtgtcat | gcattgttgt | gtgtgtgcat | gtgtgtgcat | gtgtgtgcat | gcattgttgt | 154740 |
| gcgaatgtgt | gtgtgtgcgc | acatgcattg | gtgctaagct | ctggagaaca | agccgcagg | 154800 |
| ctggggcagg | agtcgggggg | caaggggggg | acatgcataa | ttctctccac | agccctcag | 154860 |
| tgttaacagg | ctgaggttgg | ggatgcacatg | tcttagaaac | aaacttcacc | ctctgcccta | 154920 |
| aaactgcctg | ctgtgtctgc | acctcgggcc | aagctgtggg | tacttggaaa | gagggccctg | 154980 |
| ctctgtcctg | ccaaaccacg | gaagtcttga | aatccccatg | gaggggacat | gggggcaggg | 155040 |
| gactccccaa | gcagacacag | caccacacaca | agggctctaag | ggaccctttt | ctccccaac | 155100 |
| gctcagggtg | gggatggatg | tggacaagca | gctcagggct | ggggctggcc | taggttggaa | 155160 |
| gtctcagaag | acagaaggca | gcagctgccc | aaggggcccg | actatggggt | ggaggctagg | 155220 |
| ggctggaaaa | agagttttgg | aggtcagcca | aggggaagtcc | agcccttggc | tcttctctgt | 155280 |
| gtcatgggga | gccaccggga | gtataggagt | ggatgatggga | tggccccaga | gttgtgcctt | 155340 |
| tgaagatcac | ctctgggtgc | atagtaggat | ggactggggg | cagggtctgag | aggaaggcca | 155400 |
| tctcatcaga | ggggaaatcc | ctcacccacag | tccagaggaa | aacagcgggc | ctggggagag | 155460 |
| gctgctcctg | gggaggaact | actcagagcc | tgattctctc | tcgtgctaac | ccaccaggg | 155520 |
| agatgcagga | agctgggtgag | gtcagtgggg | gggcccacaa | ccatgctgga | gaggagcagg | 155580 |
| agggagacgg | gagggcacag | agaggtagag | ggcaggggact | ctgaggaccc | tcactccatgc | 155640 |
| ccacctggac | aattgaatcg | agcttctctc | agtgtggaat | gtgggctcac | tggggctatt | 155700 |
| tgctatcctt | ctctcatgtc | agactggatg | cggcccaagg | gtggggtctg | tgactccctc | 155760 |
| atcagactga | ggggggccga | aatgtggggg | tgggggaaag | agggagcttc | caaacaggaa | 155820 |
| ggttttggagg | gcgctggcac | tgaggacagg | ggatacgccc | cccatgagtc | ttcacccctg | 155880 |
| ctcagagcct | gctgccgtgg | ccctcagaga | cgagcccatc | tttggagctg | gagttgggga | 155940 |
| cagggaaccc | aggttgggaag | gaaaaagaaa | aaaggggcag | gcatgagtgt | ggggggccag | 156000 |
| accctgcaga | cccaggaca | tcacacacac | acacacactc | caatccctcc | tggctcttgt | 156060 |
| cagcctcaaa | cgaggtgcag | gcctgcaggc | cccaaccaag | gaacctgttt | acacacagac | 156120 |
| acaaagacac | aagatgcagc | aacacacaca | cgtaatacac | aaacacagtc | acacactagc | 156180 |
| gcacacacag | acacagtcag | acaatccaag | aaacacactc | acacatgaac | agacatatata | 156240 |
| gcaaagtctt | acacatacaa | agacacactt | aagcacagac | acaaacacac | caacagatgc | 156300 |
| acagacggac | acacatagac | gcacactggg | aggcgccctg | tggagactgc | tgatttgata | 156360 |
| ccatgttggc | ttggtatcca | gaattgttgg | tgcagcgcac | tccccatgcc | tgacctcatt | 156420 |
| aaggaggccc | tctggggcag | tggctctgag | aagaaaggca | gcaagcctag | ggaggtgaca | 156480 |
| ctgtagctgg | ggagcatctt | ggctggggag | agaggaggaa | ggagagcatt | tcaggggagg | 156540 |
| aaaaatagccc | ttgcaaaagg | ttggagtgga | agagcaaaag | ggatgtccca | agaactgtgc | 156600 |
| ctcaattgat | ggggctgtgg | ggcccttctc | tgcagggcag | ccttccagg | cacaaaggga | 156660 |
| ggtggactga | gaagcatctg | ggaacttctg | gggggcacat | ggggcagagg | ggagggggaa | 156720 |
| cagggcagaac | gagtcagctg | cctttgaaag | ctccacgagg | gaacctagac | aaattcatct | 156780 |
| attcatccgt | tattcatctt | ctttatgaca | ggttctctct | ctgtcaccca | agctggagtg | 156840 |
| cgggtggtgtg | atcacactta | ctgcagcctt | aaactccttg | gctcaagcaa | tcctccaccc | 156900 |
| cagcctcctg | agtacctagg | actacagacg | tgagctacta | caactggcta | atcttttaaaa | 156960 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| cttttttagag | acagggtctc | cctttgttgc | ccaggctggt | ctttgttgcc | cagccagctt | 157020 |
| caagccatcc | tcccacctca | gactcccaaa | gcgctggatt | acaggcatga | gccacggggt | 157080 |
| cccagcccgag | aaagagacct | tcacccccaca | aacccctgt | ctggaggcca | agtcccagca | 157140 |
| tcccagaaag | ggccccgggt | aggacaagaa | aaaggctgtg | tgccctggag | gagaggctgg | 157200 |
| gtgccccctg | gtgcggagtg | gtccccttgc | cgggcatggg | gcctcagttc | gtgagactcc | 157260 |
| ggccctccgt | cgctctgaact | tgagcaaaag | cagagctagg | cctagctctg | ggaccccgag | 157320 |
| gcctgcctcc | atccgcaatt | ctgcctcttt | tctggccact | gaggaccggg | ctagggaaag | 157380 |
| tgtaggggca | gtagggggtg | gaggagagag | ggttccgggt | caagggactg | gaggggtctg | 157440 |
| gggcagggaa | gtggaggggt | ctagggcggg | cccttccccc | tccgctgttg | ggccctcgct | 157500 |
| gaagggggccg | aggctgagcc | ccaaggtggg | cctgtggggg | cggggcgggg | gagccctca | 157560 |
| gcagctggac | cgcaggaggc | cgaccagggt | ctggaactcc | tcttccctcc | cctccgccgc | 157620 |
| ccctctccc | tctcccacc | ccctgggctg | ctagggggagc | gggggaggcg | caggaggggc | 157680 |
| tcaggggagg | gcccctggagc | cgggaccagg | ctgggcccct | cggcgaggcg | ccgcgcaggc | 157740 |
| aggccctgcc | gctggcctcg | cacatctggg | ccgggagcgc | gggcccagcc | ggcgccgagc | 157800 |
| gcggcgccgc | catccgctct | gggggagggg | gcggcgccgc | gggtggggccg | gccccagaa | 157860 |
| ggcggaaggc | ctcaggggccg | ggcgcacggt | gagggctgcg | accgcgcgag | cgggcacggc | 157920 |
| ccaccgagct | gggggccccg | gatcgcggcg | gctggacggg | gctggagctg | ctgggagggg | 157980 |
| ggaggttgagt | tctggggcgg | gggctgcggg | gcgcccgcag | tggagaaagg | cagggaggtt | 158040 |
| tgccgtccc | ggctgtccgc | tgagacccca | ccaaaaacct | ccagactttg | tctggtgggg | 158100 |
| acagatctgc | aagccctct | ctgcagcgtg | gctgcccgtg | cgggaaggca | cttggggcag | 158160 |
| cgaccttgggt | ctcccacttg | ggcttccggg | gaccacctcc | cctgtccctc | gcacagctgc | 158220 |
| tcagaggctt | aggctggaga | aactcctgtg | tcttaggggc | tggcaggact | ggtgtgggct | 158280 |
| ggggcctgct | ccaggggagc | aaactgaggg | accgggtgcc | taggtctgag | gtggaaacct | 158340 |
| gggagttag | tgtatgtgtg | tgctgttggg | ggtgcctcag | gggtccagg | ccctgtgttc | 158400 |
| ccgaaggcca | acagagtact | gagtgggcca | ggtgccagg | tctgaggctg | ggacctgtgc | 158460 |
| ttccggggcca | gcttgaggtc | cagggtgggt | ggggtccagg | ttcccgcctag | ggcagttggag | 158520 |
| cctatctgtg | aggttagcgt | gtgttgatga | gagagactgg | aagggagtga | tcacgaaact | 158580 |
| aggggaccctc | cggagagcag | cggcagcgca | ggaatccag | gccagagggg | taatgggggg | 158640 |
| ctgtggggcag | agcatgggag | agggctctgt | gtttgtgaac | tgctgtctcc | catgtgaggg | 158700 |
| acccaagggg | gtgcggcgag | atgtatggaa | agcaaagatg | aggacaggga | ggggccatgg | 158760 |
| gaggccattt | ctgtcactcc | ctgccccagc | gcaggatggg | gcagccacct | cactcagcca | 158820 |
| cacatggtcc | cactcttcta | agggctcctc | tcagaacag | gttggggcag | cccttggggg | 158880 |
| gcggtttctg | atctgtgctt | tgggcccctg | ttggcccagc | ctgtggcttg | gaagaaacgc | 158940 |
| ctgtggcctc | acttagcccg | ctgacggagt | cctgttccat | tacagggtct | gctgggtgct | 159000 |
| gagcagagag | cagcaggcgg | aagaggcagg | gtcctagtga | ggaggggagc | tggaatgacc | 159060 |
| tggttggggt | ctttgtctac | acaggcatcg | tgtggagccc | tgaaccagcc | aacactgcat | 159120 |
| tcccatagtg | ctcctggggg | catgagccgc | cagctgaaac | accccacaca | ggaagagatc | 159180 |
| ctccgaggat | tctgtcaggc | aaaggggttg | gcccctgcaa | ggtggctcac | gcctgtaatc | 159240 |
| ccagcacttt | gggaggatga | agtgggagga | tcgcttgagg | ccaggtggct | gaggctgcag | 159300 |
| tgagccgtga | tctgtccact | gcactccagc | ctaggcaaca | gagcaagacc | ctgtctcaaa | 159360 |
| aaaaggagcg | ggagtggggg | cagggttcgt | cctaggctat | tgtaaagcca | ggggaatact | 159420 |
| tgggattggga | gtctctgggt | gtgagttcct | gctggggcac | accctatcct | ggtccccagt | 159480 |
| gactaatcgg | cctcgcttat | tgggtagaag | ctaggggcag | ggctcagggt | cagctgtgatg | 159540 |
| actagaggtc | aatatgaagc | caagagttag | gggtcagtc | atggggcagg | ttagggatca | 159600 |
| ggctggggcc | agagtttggc | atgtgcttgt | agccaggaaa | tggagttcct | tctgtgttaa | 159660 |
| atcaaaagtc | cccactcaat | tactctcaaa | cccattaggc | ccaccagtct | gaaattagta | 159720 |
| ggcccccgtga | gcgggggtca | gcctccctcc | attttcagct | gtggactctg | cccagaccta | 159780 |
| gctttcggga | gtcgggcttg | tccctgtgtt | tgtggtcagc | actctctcta | cctgtccctg | 159840 |
| ggtgtgtgtg | ctcgcttggc | agtcagccac | ctcttggtct | acctccacct | ctcgggggat | 159900 |
| ccatgggatc | ctccttcttt | aacacaaatt | tttttttttt | tttttttggg | gacagggtct | 159960 |
| cactctgtca | cccagctggg | agtgcaagtg | tgtgatoatg | acttactgca | gctccaacct | 160020 |
| tccaggctca | agcaattctc | ccactctcagc | gtccaaagtg | gctgggactc | taggcatgca | 160080 |
| ctaccacgct | gggctaatat | ttgtattttt | cgtagagatg | gggttttgct | gtgcccagg | 160140 |
| cttgctctga | cactccactc | tcaagcaatc | gacctcccaa | agctccctga | agtgctagaa | 160200 |
| ttacaggcat | gagccaccac | acccagccct | taacacaaat | tgatttagtt | cctacctctc | 160260 |
| tccttaaaga | actagatg | actaagatgc | gattctcagt | aagaccaa | tatctaccaa | 160320 |
| gaaaattcag | aatctgcac | atgaagggaa | agaacactgc | tcaaacctgc | ccaacctcct | 160380 |
| gcttttatca | ataccttga | tatctgggact | tgttaacata | agccctctgt | gtcaccctatt | 160440 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ctacaggaaa | gggcactgag | gcacagagat | gtggcatgat | gcactcacgc | taggtgccat | 160500 |
| aataggctct | tgccatccct | gaaaccccag | cataggggca | ctcagcatat | cgcacaatga | 160560 |
| gagacggtag | ataactggac | aagcttcocga | tagctgggaa | gagtataggc | ccatgggttg | 160620 |
| gatgtggaa | ccagtgtaga | gcgaggactc | agtctttttt | tttttttttc | ttgagatggg | 160680 |
| gtctcaactct | tctcgcccag | gctggaagtgc | aatggcgcgga | tcttgggtcca | ctgcaacctc | 160740 |
| tgcttcccag | gttcaaacga | ttctcctgcc | tcagcctcct | gagtagctgg | aattacacgg | 160800 |
| gtctgcocact | acgcttggct | aattttttgt | tttttagtag | tcacaagggt | tcaccatggt | 160860 |
| ggcaggctgt | gtctcgaatt | cctgacctca | ggtgatccac | ctgcctcagc | ctctgaaagt | 160920 |
| ctgaggatta | caggagttag | ccacggcgcc | cggcccagag | cttttttttc | tctggtgggt | 160980 |
| ctagggttgt | ttctcctccag | gtacaaatagc | ctcgggtccag | aaaatccatt | cattcattca | 161040 |
| acaaagtgggt | tcccagctgc | ttagcacaga | gcctggcatg | aagaatctcg | ataagtgttt | 161100 |
| gctaagaat | cagaacaaca | tacaaataag | gctgcagct | cctccctcga | gaccttgatc | 161160 |
| cagtagtatg | ctggttaatg | gtttccaatc | agctctggaa | gaaggataaa | atcctgattt | 161220 |
| atggtgtttg | ccaattttctg | tcttgttaa | actcccagcc | atgactgatt | tgaaagtacc | 161280 |
| aatgtgatgt | cattgaatgc | agaattggga | agagatagaa | agaaatcagct | cgatgagct | 161340 |
| ggtgggaccc | agctcagcat | cccactccat | catcctgagc | tgggctctgg | gctcctgccc | 161400 |
| ctggctccta | agaacccctcc | cctcccccctc | tctgtgcttt | ccctttttctc | ctcaacagcc | 161460 |
| agctgctata | tgcatgcccc | caaccagcac | ccatccttct | cttgattctg | attcctgagc | 161520 |
| atccaactcc | aatggtacag | aggacagcaa | gagtgctctg | aaatctccat | ggtgctggcc | 161580 |
| atgcaggaga | ctctgggaca | tacaagagtg | tgtgggagtc | cagggtgctg | tcacaattgg | 161640 |
| ggctgggttg | tgtgaactgt | gtgagcaagg | gtgtgtgcat | gtgtctatgg | gcagtgcaga | 161700 |
| gccagctctg | tgcttgcctg | tgttctatag | caaaactttc | agattggggc | gctgggtgaca | 161760 |
| gagaaggaa | gcatttggccc | tctcctttgt | ttagaggagc | ttagaggagc | agccaggccc | 161820 |
| accaatgaga | acatagtgtc | agtggtctgc | agactggctg | agagacactg | ttgagcaaca | 161880 |
| cttctcaaac | ttcaatgtgc | acctgaatca | cctgggaaatc | ttgttaaaat | agggccagagc | 161940 |
| cggctggccta | cgctctgtaat | cccgagcact | tgggaggctg | agacggcgcg | atccactgag | 162000 |
| gccaggagtt | cgagacagc | ctggccaata | tggtgaaacc | ccgactctac | taaaaataca | 162060 |
| aaaattagcc | aggcattttt | aggtggcatt | ttaggcaaat | taggtggggc | atgcctgtgg | 162120 |
| cccagctac | ttgggagct | gaggagagat | aatcacttga | accagggaga | tggagattgc | 162180 |
| agtgagctga | gaactgagat | cacgcgactg | catctcagcc | tgggggacag | agttagagctc | 162240 |
| tgctcaaaa | ataaatat | ataataataa | ataataataa | ataataataa | ataggaata | 162300 |
| catgttctga | ttcagcaggt | ctgggtctggg | gcttagaggt | ctgcatttct | aaaacagcttc | 162360 |
| caggtgatac | caatgtctgt | gatcctcagc | ccacactttg | agtaacaaag | ctgtgtgattg | 162420 |
| tttgaatat | cctgcctata | cctcgtcccc | agaacttcat | agacaggatc | gagccctcgc | 162480 |
| catcttagaa | aataaatatt | cttctctacc | tcctctcagc | tttagcgccc | tcacactact | 162540 |
| gccctcacat | gtatttctga | tttcagccaa | gcttttcaga | aatcttagct | caatatgatt | 162600 |
| gcgtcactac | actccagcct | gggcaacaga | gtgaaactcc | atcttaagaa | aagaagttct | 162660 |
| agctcatgtc | acctcttgct | cactcctcac | ctaagcttac | ctcccttggg | ccactaaaac | 162720 |
| agtggttcca | gtctcatctg | gggacctttt | acagggatct | agctttacc | atctcttggt | 162780 |
| acaactctca | tgagaccatg | gtctcctggg | gtctcctggg | ccggtatata | ttcctctggt | 162840 |
| cagcctttaa | gttaggttct | tcctcgtgac | agagtgcac | cctatctcaa | aacaaacaaa | 162900 |
| caacgaagcg | caggcacggt | ggctcatgcc | tgtaatccca | gcactttgag | aggccgaggt | 162960 |
| ggggcgatca | cctgaggtca | ggagttcgag | aacagcctgg | ccaacttggc | gaaaccccg | 163020 |
| ctctactaaa | aatacaaaaa | ttagctgggc | atggaggcac | tcacctgtaa | tccagctac | 163080 |
| tgaggaggct | gaggtaggag | acctgggagg | cagagggtgc | agtgagcttc | agtgagccaa | 163140 |
| gatcgtgcca | ctgcattcca | gcctgggcaa | cagagtgaga | ccctgtctca | gaaaaaagaa | 163200 |
| aagacaaatt | gagtcagctc | cccccccac | cctgccacct | caaacccttc | aagatttgct | 163260 |
| gttgcccttc | atataaaa | taagggtgac | agggcctgct | agggcctgct | tgagggttaa | 163320 |
| gccagcccca | gccctctgtc | cagacttacc | atgtcctctt | ccacccccgc | attatccaac | 163380 |
| tttagccttc | ctagactcca | gcaacgcttg | tcatttcaaa | atcacttgcc | tgattatctc | 163440 |
| ctcatcaga | cogtttggtt | ctagggcagg | gacagtgctc | ctctgggtca | ctatctccca | 163500 |
| gcacctggca | cagagcaggc | accagaaaac | atttagaagg | atcctgattc | caagaaagta | 163560 |
| aagccatttt | ttgacacaag | ctgggaactg | tgggaactta | gatgagattc | aaggaattac | 163620 |
| tgttaaatttc | attagacatg | agaattggtg | tgccatgaca | ttgtttaaaa | cgtctttatc | 163680 |
| tgttggagag | atgtatataa | gagttcacgg | gtgaaacac | atgagtggct | ggatttgctt | 163740 |
| caaaaggactc | cagccaaaac | taaacaaaata | aaataataa | gcattcagat | gagggacaaa | 163800 |
| agggacaaat | aaatccact | ctgcttgctc | gttctctgtg | acagctacag | ggcctgcaag | 163860 |
| gatgttggtc | gtcccggaga | tgggcctgca | ggggctgtac | atcggtaaga | acccccacca | 163920 |

| | | | | | | |
|-------------|------------|-------------|-------------|-------------|-------------|--------|
| ttctgcctcg | acccatcacc | cactcaccgc | agcaggcact | gagacccacc | tacagggtgt | 163980 |
| gcaatgtacc | tctcccacag | agcttcccaa | cagctcttct | gtgagtgagt | gtgtgtgtgt | 164040 |
| gtgtgcattg | gtgcaggcca | ggctctgtgt | tactggaccc | cctctgtgtg | gcagaaaagg | 164100 |
| tctgtgtgca | coggtctgtg | tatgtgagta | tgtgtgcaca | caggactggg | ctattccacg | 164160 |
| gacctcatg | tgtatgatag | tgtgggctga | gtgcgtgtgt | ttgtgtgtgc | cagaaggggtc | 164220 |
| tgtgcacaca | ggagtgtatg | catgtgcgtg | tgttttgttc | aggtctctgc | ccaggaccgc | 164280 |
| tgtgtgagtg | tgtcagtgag | tgtgtgtctc | tgcgggcagg | tctctgtgca | cgggatctgt | 164340 |
| gtacgggagt | atagtctgca | gggtgggcatg | catacagaga | ttctgtcggt | tgcattggagg | 164400 |
| ctctgtgtgc | acaggyagt | tgaatgttc | atgtgttttg | tccaggccat | gtgcacagga | 164460 |
| gcccgtgtga | gtgcatcagt | gagtggtgtg | ctctgcgggc | aggtctctgt | gcacaggatc | 164520 |
| tgtgtacggg | agtatagtct | gcagggtgggc | acgcatgcag | agagttctgc | gtgtgcatgt | 164580 |
| gcgtctggcc | atagcctggg | ctgggagcag | gcttgggttc | tccgcagggg | tgggaggccg | 164640 |
| gaggcaagag | ggacactcct | caggaaagtc | attcacaag | acctgggtggg | ctgggcccct | 164700 |
| ccttctcgcc | tcccttccca | agctcgagcc | acagcaactt | ggcctctcgg | gctgcatgtg | 164760 |
| gtggcagggc | tggagcctgc | cctggctgcc | taggagaggg | gagcccttga | aatgtggcca | 164820 |
| cttttccctc | gcccctccag | cacctcagac | tctggccctg | tccaggcgct | tcccagagtc | 164880 |
| ccccctcatg | ctggcccctc | gcctgtgttc | ccaggttacc | cccaaggcca | gggagagtgg | 164940 |
| agacctgctc | ggcagcagct | gacagagactg | tggctgaggg | agcacagggg | ctgctgggct | 165000 |
| cagcctaagg | cgagccagac | ctgcctgatt | tgtccatgga | gcttctacc | agatctcact | 165060 |
| ctccctccct | agccagctc | ccatggctgt | ccatggatat | caaggctgcc | ccagttactc | 165120 |
| tcacatgccc | cttagcaatg | cccagtggtg | tgggtggcag | taccaggccg | gcaggcccctc | 165180 |
| ctgccccata | agccgctcga | tctctggggc | gaggcctggg | gctgtcctgc | agggctactc | 165240 |
| tactccctcc | caccaaagg | ccagctcagt | cccaggcgtg | caagaggggc | ttgatggggc | 165300 |
| gtctgcccct | gaaatgctgg | ctcagggtgt | gtgtatgtgt | gtgtgtgggg | ggtgggaatc | 165360 |
| agacttctca | atgttgttga | caccccagtg | tccaggactgg | gctatctcta | tgggtgtggc | 165420 |
| aggaggtgct | ccccctctc | gatcctctc | aagcagccag | ctgggctctg | ctgtggccct | 165480 |
| tgtgggcatg | ggagctgca | agtttccctc | agtcttgatt | agaaccacgt | ggcccaactt | 165540 |
| taggggtggg | gggagcaacc | ccactctcta | gccagttagc | tctgtagctc | caccttttga | 165600 |
| aggaggagg | cttctctctc | tctactccc | tcccaggagc | tactctgccc | tatctccctc | 165660 |
| gccccttccat | ctctggactg | agttctccct | ccacacaagt | ctttctgtgc | tctctccctc | 165720 |
| cccacactgt | gtgtctgagg | ggccttagtg | gggagacagt | gagctctctg | ggaaagcatg | 165780 |
| ggatcatctg | tagggaaaag | tgtgctctga | cagcagaccc | tccagctcct | ccccgtcgct | 165840 |
| atgtgtgtct | tctctacacc | cgcgagctct | cctccccagc | agagagccct | gtcccttagc | 165900 |
| ccagcaaggga | actgttgaga | gccttgggtg | gggcccgcgt | ggcaccctca | gcataacctc | 165960 |
| ccctgcttgt | gccaggatga | actgcccgtc | cttctctgca | ggctccagcc | cggagcggtc | 166020 |
| cccagtgctc | agcccacccg | gctccccgag | gacccaggaa | agctgcggga | ttgccccctc | 166080 |
| cacacctctg | cagtctccag | taagcccaga | gcagggaacca | ggtgtgggtg | gacctggctg | 166140 |
| gtgtggacag | ggtcgtgcgt | ggcaaaagtc | tgcaggggcc | tacgttagga | aggaggagg | 166200 |
| gatgggggta | gcaccatgcc | tcttgtcccc | gtacactcca | gtcatgtgcc | cccagaaagg | 166260 |
| attggccttg | ggtgacctg | gactataaag | ggtaacaatt | ctactgagaa | ggcgagcccc | 166320 |
| atgtaactaa | ctccccaccc | ccccacacag | gaggcaccca | cggccattga | agctggcctg | 166380 |
| gctgaggcaa | gtgttcccc | gagcatgtgg | cttttctagg | tccccagagc | actgggccc | 166440 |
| ttgtggacca | cagccccact | taatccaagt | catatctgaa | ggagaggaga | tgttgtgcca | 166500 |
| aaaaaaaaaa | aaaatagtct | ggaaaccatt | agctttgggt | gggcatgggt | gctcatacct | 166560 |
| gtaatcccag | cactttggga | ggccgaggtg | ggcagatg | tttgagctca | ggagttcaag | 166620 |
| aacagcctga | gcaacacagc | gaaaccccat | ctctacaaaa | aatacaacaa | ttagctgggc | 166680 |
| gtgggtgtgc | acacctgtaa | tcccagctac | tggggagggt | gaggggtggg | aattgtctga | 166740 |
| acctgggaag | cggaggttgc | gatcgcaacca | ctgcatttgc | gcttaggtga | gcctaggtga | 166800 |
| cagagtggga | ccctgtctca | aataaaaaag | aaactactag | cttgaagtat | ccaggacacc | 166860 |
| acctccccac | ccccagctc | cctgggtcac | cccacctgga | tgaccttgcc | ctggtgatcc | 166920 |
| agcttaggga | cccccttatt | cttaggggag | tccagccctg | gcataaggaga | aggcacacac | 166980 |
| ttctctgggg | acatgcacgt | gcccctctct | ctctgcccac | gaatcacaga | gtctggtttg | 167040 |
| tgtcactatg | gccacaagaa | gcaagaggat | aaaatatata | cactggtcca | tgtgctgtaa | 167100 |
| aaactgctgg | aggagcaaaa | ttgctcccag | ctggggccag | actaatggct | gcctatgaag | 167160 |
| atggtgtgag | gtgggggtgg | ggctaaaggca | gtccagagg | aggggtgggg | gaggctggaa | 167220 |
| gcagggaagc | aggagcttga | gccaacacct | gcctggggct | acctgagaga | cagctccaag | 167280 |
| gctcagcctg | gagcctggga | gggcaaggga | gcccagagc | gtgggcaggt | gggggtgggc | 167340 |
| aggtccagg | ctggtctgga | ccacagtcag | tggaggggag | tgtcttcccc | atgcagagga | 167400 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| agcattgggg | ctgtggggga | tgggggtgtc | cctgctgggt | gotttgcta | tctaagtctt | 167460 |
| gcacttgca | aaaccocagg | tccgagcccc | ccagcaggcc | tccttctctg | tggttggtggc | 167520 |
| cattgacttc | ggcacccagct | ctagctggcta | tgtcttccagc | tttgccagtg | acctgagggc | 167580 |
| catccacatg | atgaggttgag | gtcggtctggg | ctgagagagt | gaggtggggga | gggtgggggag | 167640 |
| ttctctatac | cttggtccca | aaagtactgt | caccgagaca | tgggggtctcc | ttggaggccc | 167700 |
| tgcacccccc | agtcctggggc | tcccactctgg | ggtaaaagt | ggaggatggg | ccccagggct | 167760 |
| gggtcccttgg | cctcaactgg | agcagctccc | aacctctctg | aggaaaagtc | agactttggt | 167820 |
| tttaatttttg | aattttttttg | agatataaca | aacataacaa | taaaatatgc | aaatgttttt | 167880 |
| tattttctct | tttttttttt | tgaggcggag | tctcactctg | tcgcctaggc | tggagtgcac | 167940 |
| tggcacgata | ttagctcatt | gcaacctcca | cctcteggg | tcaggcgatt | ctcctgcttc | 168000 |
| agctcccca | gtagctggga | atacaggtgt | gcgccactat | gcttggctaa | cttttgttgg | 168060 |
| gttttttgtt | tttttttgag | accggagtctc | actctgttgt | ccaggctgga | gtgcagtgc | 168120 |
| gcaatctcag | ctcactgcaa | gctctgcctc | ctgggttcat | gcatattctc | tgccctcagc | 168180 |
| tcttgagtag | ctgggactac | agggcgaccgc | caccacaccc | agctaatttt | tttttttttt | 168240 |
| tgtattttta | gtagagcgg | gttttacocg | tgttagccag | gctgggtcca | aattcctgac | 168300 |
| ttcgggtgat | ccgccagcct | cggcctccca | gagtgtctagg | attacagatg | tgagccacca | 168360 |
| cacctggcct | tttttttttt | ttcttttgaga | cagggtctcg | ctctgtcaca | caagctggag | 168420 |
| tgcagtggtg | tgtatcatag | ctcactgcagc | ctcaaacctc | tggggtccaag | caatctccc | 168480 |
| acttcagcct | cctgagtagc | taggactaca | gacagcacca | ccacacctgg | ctaatttaaa | 168540 |
| aaaaaaaaa | attttttttt | ggagagacag | ggtctcactg | tgttgccagg | gctgggtcttg | 168600 |
| aattcctggc | ctcaagtgt | tatcccactt | tagcatocca | aaaactgttg | aattgcaagg | 168660 |
| gtgagccacc | atgcccagcc | cacaaatatt | tttctaagct | tctagacaaa | cacacacata | 168720 |
| tacaacata | tatgtgtttt | tttaagtaaa | agtgcagata | acaaactgct | ttgttttaaa | 168780 |
| gcttcaacac | tatttttgtag | atatttcatg | tcagcacata | cgaagctacc | actttctttg | 168840 |
| aacggcctag | tattccatag | tacagatgtg | tcataattta | tcacttcccc | tattgcagct | 168900 |
| tttagggggg | ttttttctct | tcactatttt | tcactattac | aaaaagtgc | gcagtaatat | 168960 |
| ccttctttga | aaacttgtgc | cgcacatctcc | atggggtaga | ttcctagagg | tgattcctag | 169020 |
| tccaacatc | tcaaatcttt | tttcttttcc | tttttttagta | aaaaagaaaa | gaaaaaataa | 169080 |
| taaaaaaga | aataaaaaata | aaaaacaagaa | aatgaaggtt | ctaaaggctg | agttgacaag | 169140 |
| tcacacagtt | gttttgaaag | acatgtttca | aaaatccatt | caaagtgttg | gttttatgtg | 169200 |
| ggacctcaag | gaccttggga | cctcaagacc | ccagtcctct | gaatatgtcc | tgaggggctg | 169260 |
| gcaggcagta | ggtgaagcta | caggagacct | caaagctctc | aaccagagct | ggatggggag | 169320 |
| gtaaggaggg | cttcccagag | cctggagact | ggatgggaag | gtgaattggag | acatagacag | 169380 |
| acaggtaact | gctcccaaaa | tacactggac | ttgagcttca | tgagactgga | ctgtccttcc | 169440 |
| ttccacatca | caggaaacca | caaaggctct | gagcacttac | agccctctag | gagctgcagg | 169500 |
| gcaccccaga | tgtcttgtct | atcaacacac | atggctgagg | ccatgctggg | gctaaggaca | 169560 |
| catccaggaa | cactgtctct | gtcctgttga | ggtgcagtc | acaggaaata | actgggggac | 169620 |
| attccagcct | cctcccagc | tctctcttcc | caccagacac | caaggcccta | acatctctag | 169680 |
| aatacccato | tacttctctc | catggccact | gatcgtgccc | aagttcagcc | ccgacctctc | 169740 |
| ctcactggga | tgggtgcaat | atcctctccc | ttggtttctc | ttctccatgt | tccttctcct | 169800 |
| gcattctgtt | gctccatggg | agccagtgac | ctttctttct | aaaggcaaat | gtggctcatc | 169860 |
| ctcttggtca | ctttaaatcc | acgaagacag | aattctgaat | ggccttctgt | gctcttcatg | 169920 |
| atctacactt | agcatctctg | tgttttctcc | tgtctgcctat | ggagccgacc | catagctagt | 169980 |
| tctcctgtct | ctttcacatc | tgttaagtttt | tgcacatgct | gttccctctg | cctggattcc | 170040 |
| caccacaaga | gggaagggttg | cactgactgc | cctgtgctcc | cgcaggaaat | gggaaggccg | 170100 |
| agacccgggc | gtggcccacc | agaagacccc | gacctgcctg | ctgctgactc | cggaggccgc | 170160 |
| cttccacagc | tttggtctaca | cgcgccgcga | ttactaacat | gacctggacc | ccgaagaggc | 170220 |
| gcgggactgg | ctcactcttg | agaagttcaa | gatgaagatc | ccacgcgcc | cgtgtgactca | 170280 |
| cagggtccca | gacaggggag | cggggccagc | atggaaaagg | gcagggtcaa | tgggggtggg | 170340 |
| tgggacaaaa | ccaaaaactg | tgaggaccgg | cccgatggag | ctgtggtcga | gagggggcgg | 170400 |
| ggctaaaagg | agacgtcggga | ctccggtgtg | ggcggagctc | acgaattgag | tggaggcggg | 170460 |
| gctaattgtg | gtggggctaa | tagtgaagct | ggggttgccg | gaggggtggg | gctaaggaga | 170520 |
| ggggctgggg | cagagctaat | gtcacatggg | gcaagagtgg | gacggttgga | aagaggaggg | 170580 |
| gaagctccag | gaaaacgggtg | gattttaaga | gcgaggtcgt | cagggaataga | gtgccaaagt | 170640 |
| gaggcctcct | gcagagccgc | cctgtgtccc | tgccaggatc | tcacacttga | gaccacagta | 170700 |
| gaggcgatga | atggaagagc | gatgcocgc | ctggagggtg | tcgcccactg | cctgcgcctc | 170760 |
| ttcaggggagc | acgcccttca | gggtgcgctg | ggccccacct | ctgcgcagctg | tggcaggggc | 170820 |
| ccctctattt | ccctctatcc | gaaacgcgtc | ccccatcccg | tcocccgacat | tggatgggta | 170880 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| gccaccgccg | gagctcagag | gtcatcttct | ccagtaccct | cctccctttt | tgtctggtag | 170940 |
| agcctgcacc | aagccatact | gatgggaggg | gggcccattc | ttccagctct | gctgggaagt | 171000 |
| ccttcctgtg | atttgattag | tacctccagt | tccgcagagg | gctggaagacc | accctccctc | 171060 |
| caagccagct | ttcctctcac | tgcccccctc | tgtaccagga | gctgagggag | cagagcccat | 171120 |
| cgctgccaga | gaaggacact | gtgcgtctgg | tgttgacggt | gcctgccatc | tggaaaacgc | 171180 |
| cagccaaagca | gttcatgctg | gaggctgcct | acctgggtgag | gacgtgcagg | cggggcccgag | 171240 |
| aaacatgctc | aggaagggcc | aggctgtgcc | ccatgcttgc | atgcacccca | ccacccttga | 171300 |
| gaccacagag | tcatgttgga | aagaacttca | gcctgctcct | gatgggaagt | tgtagagtgt | 171360 |
| ctcccccagc | aagaggggagt | gggctctgca | gaacaggsga | caagagggaca | aaagacacag | 171420 |
| cccaggccag | tgtagacagt | gccttaagtc | agggttccct | aaagcagagc | ccaagatgga | 171480 |
| gagtcctttt | tttttttttt | tttttttttg | agacagtctt | gctctgtcgc | ccaggctgga | 171540 |
| gcgcagtggc | gtgatctcgg | ctcactgcaa | gctctgcctc | ccagggtcac | gccattctcc | 171600 |
| tgctcagcc | tcccagtag | ctgggactac | agggtgccgc | caccacgccc | ggctaatttt | 171660 |
| ttgtattttt | agttagagacg | gggtttctatc | gtgttagcca | ggatgggttc | aatctccctga | 171720 |
| cctgtgtatc | cgcccgcctc | ggcctcccaa | agtgtcggga | ttacaggcgt | gagccaccgc | 171780 |
| gcccgcccaa | gatggagagt | cttatatgtg | agggtctattg | aagaaaggttt | ctcaggagaa | 171840 |
| agggcgaggga | gggaggcggg | acaggagaag | gagctatgaa | ggatgtggct | tttgttggag | 171900 |
| tctagcctca | gggttagcct | cagctagatc | ccatggggag | ctgcagaggg | agaactgtag | 171960 |
| cacagagtgt | ggggcggtct | cttgacatc | cgtatcggtc | tgtcactggt | tctggggaaa | 172020 |
| tgggagggaa | acctctcaa | gtgaggccat | tcccattcag | ctgagggcag | ttatccagag | 172080 |
| gaggtagcag | ctgtgagcca | atagagacca | acactcacag | tggcaggcag | ggcaccagga | 172140 |
| acattcactt | cacttggcat | cagttatttg | agggacaccc | ccaccaccgc | ccacccttgt | 172200 |
| tggttcttga | gcaggggaaa | taggactatg | aaataaaggg | aggcgattct | agccaatgca | 172260 |
| tacaggtctt | cctaccactg | agatgctcaa | agggtctagag | caaggagttg | gggggggctc | 172320 |
| ccgtggagac | agagagggca | tagcctcaag | actcattgct | tgggtctaac | ccttagagctc | 172380 |
| tggatacaga | ccctaagctg | tccaaaaggg | atctcaagcc | tgcctgtgga | caccctggag | 172440 |
| ggacccccaa | tctcccbata | cctatggggc | aggcatgacc | cagggccagg | tctgtaaatg | 172500 |
| tctacacttt | tcactggaag | cagcctaacc | tgagcacagt | ggacatccct | gtgcaaaagt | 172560 |
| caggtaggaa | gcaaggctgt | ggcagccag | aacaaacacc | tgtctcccca | ggttcccttt | 172620 |
| gcctattttc | ccaccctctc | tctcacgctg | ttccctgggt | gctcaggaaa | gtctggccct | 172680 |
| gggttaaaagt | acctctcaac | ctctccctatc | ctctccctca | tctggtcaga | tcttggggaa | 172740 |
| ccctacaaaa | ccataatata | cttcccacag | ggatccctca | ccctgaagga | aataactcca | 172800 |
| gatctcaagt | gtttttctcc | accgaaactg | cctgtcttgg | tgcctggaaat | ttcagcttcc | 172860 |
| tcctctctgt | atcagattcc | agttctcctg | gaagtggggg | ggcctatggc | ggcctatggc | 172920 |
| taccgtctga | caccagccta | gtggagaaag | agggttcaat | gagctaaaag | ggctcccccac | 172980 |
| cacctctctc | gagccatcac | acacccccac | attgtgctag | agtccttgcc | aatacccccc | 173040 |
| gccagcgctc | agctggccgc | ttgtacctgt | gaagggggac | cttgtctgcag | cccgcctatc | 173100 |
| ttgggaaatt | tgaaggggca | gatcccaggg | gttctaggtc | tgcattctct | ctagagtctt | 173160 |
| ttctctctgg | gcaacttctt | ggcatatttg | gccaggggca | ttccctcccc | agcctgggtc | 173220 |
| cagctgggct | atgaggcagc | tccaaatttg | tgcagcacag | aggggcccga | gaggccctaac | 173280 |
| atgggtgggg | tgtgatggag | gaggaggttg | gagcccccac | ggccgcgggt | ccccacctca | 173340 |
| cagtgcctac | ttagggttga | cagccccacg | tgtgcctga | ccctgcccac | caccactccc | 173400 |
| caggctggac | tagtgtcccg | agagaaatga | gagcagctac | tcctgcctct | ggagcccag | 173460 |
| ggccgctcgg | tatactcccg | caagctgcgc | ctgcaccagc | tctctggaet | gagtgccggg | 173520 |
| gcccaggtg | gtgggcgctc | gggtgagcgc | cgctccatcg | actccagctt | ccgtcaagggt | 173580 |
| gagctgcccc | cggggacacc | accacccctc | ggagggtcag | agggtcactg | aagccagaag | 173640 |
| ctcagccatg | tctagtatga | aggggagagg | gtaccacccc | tggaggagcc | caattttgagc | 173700 |
| aggcagaaac | atggctgggt | gagcctgtct | gaggaaagga | ggcacagccc | tgcccaaagg | 173760 |
| caacctctgc | tcagggttgg | acatgcccta | cctaggggcag | cccaatctca | gggtggagat | 173820 |
| ttcaccttgc | cctaggggag | caaagtctga | ggagggcagg | actccgcccc | gccttgagtc | 173880 |
| tgaagggttg | gggagacaca | gactacccca | ggaggaaact | gtcggggtac | tgcacaaagga | 173940 |
| ggaagtgcag | ccactgcata | agctgtcccc | ttcactcccc | tcatctcccc | cgacacacat | 174000 |
| cagccccaa | cctctgctgc | caggaccagg | caccaggccc | aaaggacagt | atggctcatcc | 174060 |
| ttccaaatgc | catctcccag | ggcaggggaga | agccctaagt | cctgagtcct | ctctgagact | 174120 |
| ccaaagacct | acctgcctcc | gctcctccaa | acccctctaa | ccctgatttt | gccatgacct | 174180 |
| gagacctctc | gcgttaaaag | aaggccctgt | gtccataaat | atttccccac | agctgttggga | 174240 |
| tacagggtgg | gagtttgggg | ttcaggattg | cctctcccca | gtcaggagca | ggttggagtt | 174300 |
| tcaggagcac | tggtgtctcc | cagtgcctcat | ggaggctcctg | ggcaggagga | tgggagttga | 174360 |

| | | | | | | |
|-------------|-------------|-------------|--------------|-------------|-------------|--------|
| acgceatagc | tggagcacct | cettctaate | tcaactccctg | ctgtctccctg | acccccagct | 174420 |
| cgggagcagc | tgcgaaggtc | ccgcacacagc | cgcacgttcc | tgggtggagtc | agggcgtagga | 174480 |
| gagctgtggg | cagagatgca | agcaggttagg | gggaaaggggg | gacggaggtgt | tatccttggc | 174540 |
| ccctaccggg | caccatatac | tgatggggggg | aagggcacatgt | ttgcaaaagcc | cgtctcttcc | 174600 |
| tcctccattc | gctgtaccca | acctggccgt | cccctcacag | tcaccgcgac | ccccacccca | 174660 |
| ctcacagcgg | cgcctctaac | tcccactctct | ccagggggatt | ctccgcggag | gctcgggtgtg | 174720 |
| agttgcagag | ccctctggaac | catttctcccc | ccacaccctg | cgcccattatg | tgggtgtctg | 174780 |
| aggttcaagc | acctgaagcc | cctcacgtcc | ctcccccgac | cctgcagaca | ggccttggga | 174840 |
| cccggggcag | ggctggaggc | tgggcgaggc | tggaggggggc | gcaggggctga | gggtgcgagg | 174900 |
| ccgcccacga | gtgtgtgccc | gcgctcgccg | ccgcaggaga | ccgctacgtg | gtggccgact | 174960 |
| gcggcgagg | caccgtggac | ctgacggtgc | accagctgga | gcagcccatc | ggcaccctca | 175020 |
| aggagctcta | caaggcatct | ggtgaagtgc | caggcgggcgc | cccggtaccc | agcgcgacce | 175080 |
| gggtctccgc | cccgcactg | ccccctggcg | gcccggcgag | cgcgtacgcc | ctcttcgccc | 175140 |
| cctgtctcac | cccagggggc | ccttatggcg | cgggtggcgct | ggacctggcc | ttcgagcagc | 175200 |
| tgtctgtccg | catcttcggc | gaggacttca | tcgccacctt | caaaaggcga | cggccggcag | 175260 |
| cctgggtaga | ttctgaccatc | gccttcgagg | ctcgcaagcg | cactgctggc | ccacaccgtg | 175320 |
| caggggcgct | caacatctcg | ctgcctctct | ccttcattga | cttctaccgc | aagcagcggg | 175380 |
| gcacacacgt | ggagcagcgt | ctcgcgagg | gcaggtgggt | cctgagcccg | cgggctcagg | 175440 |
| cagggtttgc | cgacccggga | atgaccggtc | actggagggt | cccgggcccc | aaggaaacggt | 175500 |
| gggggtctgc | ctgattcatc | ccacatatac | actaagccag | caggggcgctg | gggtggggcg | 175560 |
| gcggggagcg | gcgagtgagt | gcccagccc | agcaggctcc | acccacggaa | tcocgagccc | 175620 |
| gaactggggc | aagacagaga | atcatagcgg | ggaggcggca | atgcctatct | cctcccagcc | 175680 |
| ttctctacac | ccccacccc | ggcctgcggg | gcccattgctc | ctcggtttcc | gtgacccaaa | 175740 |
| gcaaggggag | gcccctccca | ggacctcgta | cctggaacct | ggagcaggct | gcgaactaaa | 175800 |
| tcctctgagt | gagtaggggt | gagataaagg | actaaacatcc | cgcagggtcca | gtcctccaga | 175860 |
| caccacgtgc | agtcggtgcc | caggcacttc | tgcctggagg | cagaggtaga | gaataaggac | 175920 |
| cacggacccc | aaactggggc | aagcagctgg | gcctgaccg | atggatattt | gcccctttca | 175980 |
| ccaccacacg | cgtgaacttc | gtgaaaggt | cctcacaggg | gatgtccaga | atgtcttgtg | 176040 |
| aagccatgaa | cagcgtcttt | cagccacccg | tcagcgggat | catccagcac | ataggtgagc | 176100 |
| acctgagctt | ggtccccac | ccgcccctac | atgaacaaac | agatgcagaa | taattccccc | 176160 |
| ctatcagtgc | ctgatatact | ccacacatcc | atacactgtg | atgagaccta | gaatcatcta | 176220 |
| gaacacctgc | gggatgaaat | gcagtggtga | ttaagagcta | gagggttgat | atgtagcttt | 176280 |
| gccaaaggcaa | aaaacttctg | gtgcctcagt | ttcccattt | ataaaatggg | gtgatagtat | 176340 |
| tgggtttcta | aaacgttatc | acagggataa | aatgagctga | agtcactaga | gtgagcacia | 176400 |
| tgtcttgac | acaatgtcta | ggtgtttaat | acgtgtaaaa | gcataatcct | tatctctcgt | 176460 |
| cctccacgtc | gtggtgggag | agaagtgggg | agcgtgagtg | ttggggaggc | gaagccctcg | 176520 |
| aggactcccg | ttagctctca | aagaaagtgc | tcaaattggct | actttctagt | gcacggtag | 176580 |
| gtacaggcta | gggaggggag | gcgcgggtgg | ccgcctagtg | gtggcctcag | tggctctctc | 176640 |
| tcccccgcc | cttctctctc | gccccttca | cccgcgtccc | cccgctctgt | ccgcgagagg | 176700 |
| ccctgctgag | acggccggag | gtgcagggtg | tgaagctgct | gttctctagt | ggcggtctcg | 176760 |
| ccgagtcagc | ggtgctgcag | cacgcgggtc | aggcgcgct | gggcgccccg | ggtctgcgtg | 176820 |
| tcgtggtccc | gcacgacgtg | ggcctcacca | tcctcaaagg | cgcggtgctg | ttcggccagg | 176880 |
| cgcggcggtg | ggtgcgggtc | gcgcgctgc | cgtcaccta | tggcggtggc | gtgctcaacc | 176940 |
| gctttgtgcc | tgggcgcac | ccgcgcgaaa | agctgctggt | tcgcgacggc | cgcgctggtg | 177000 |
| gcaccgacgt | cttcgagcgc | ctcgtggccg | ccgagcagtc | ggtggccctg | ggcgaggagg | 177060 |
| tgcggcgag | ctactgccc | gcgcgtccc | gccagcggcg | cgtactcatc | aaacctgtact | 177120 |
| gctgcgcggc | agaggtatg | cgcttcatca | ccgaccccg | cgtgcgcaaa | ctgtggcgcg | 177180 |
| tcagctcaga | gcttgagccc | gcgactcgc | gccaggacac | gcgcggcgcg | cctcccggcc | 177240 |
| gccgcgagat | ccgcgcgcgc | atgcagtttg | gcgacaccga | aattaaaggtc | accgcgctcg | 177300 |
| acgtcagcac | caatcgctcc | gtgcgcggt | ccatcgactt | tcctttccaa | tgaggcgcg | 177360 |
| ccggcgcggt | gccagcgcg | ctgcgcgcgc | cccgcctct | ttcgttccag | ggcgctcgcg | 177420 |
| agcgggttgg | ggcgggggaa | acgatagttc | tgcagttctg | gcctttccac | gcctccagc | 177480 |
| cccgggggag | ataaggtcat | gggagagtg | gtgggggacac | acccagagac | tggcctttggg | 177540 |
| atctggcact | gttcgcgtga | gttcagagct | gaagggaccc | gccagggact | gaacgggtga | 177600 |
| gagaagaggt | ttgcaagaca | gagcgcgag | cccgcgaagg | ggcatgtgac | cccgaaggaa | 177660 |
| gaacgcaca | gaagagctct | ggtctgaact | tggccgagta | gggtggggg | tgggtaggca | 177720 |
| ggaggagccg | caggaggaag | gaggttgtgc | agggctctga | ctcgagggc | tgaagttcac | 177780 |
| tcacgcagcc | actcagcccc | aaccgggagc | caggcagaaa | aacctgtg | cgtaggaaag | 177840 |

| | | | | | | |
|-------------|-------------|-------------|-------------|--------------|-------------|--------|
| tgactggaag | tggactccag | agggaacaggt | gtgggtggcac | agtcctcgtgtg | tgggtcgtgac | 177900 |
| cacccaata | tgactgtgaa | ttgtggaaag | ggcagtagat | ctctaattgtg | gaggtgggaa | 177960 |
| cattattgtg | gtggaggcaa | ttatgagggg | agcatttctt | tcgagacaaa | acacccgtct | 178020 |
| gggaaggccc | caaggtcagc | ttatgaagga | ccccacttgc | acccccccc | agccatggaa | 178080 |
| gagcagctgg | agggtcggatg | gggaggccag | agggagcaat | gaggggtcgt | cccagctctg | 178140 |
| ctattgactc | ggatgcctt | taggacattc | tcttaccgct | catgggcctc | agtttctctaa | 178200 |
| agtggtgaat | gtcaggacct | ttccctctaac | tggcatgcaa | tgccccacc | tgccctgagag | 178260 |
| ccctgagsgt | acaataaaac | atttatgtct | aaggggaagc | cacagctctg | tgatatggcg | 178320 |
| tggagaccct | aatagtgga | ggaatgcaag | ggttcccggt | gctagagaga | gaaggggaga | 178380 |
| agctttcagc | tgtgcatatg | gaactgacca | gaagggggtg | ctgctgtctc | ccatcaagca | 178440 |
| tcccaaaaca | ctccactgct | taagacctct | ctggcctaca | catgaggtcc | ctctctcctc | 178500 |
| attcaaat | attgtcttgg | aagccagctt | ctggcctaaa | atgccaccac | ctgtgcatac | 178560 |
| ctcttgtggg | gctaggtgct | ataataccac | gcggtgcccc | tgctctctga | gtgagtctac | 178620 |
| ccaagctctt | ccctggccca | tctgcaaaag | agtaggcatt | acccccacc | cagagaacaa | 178680 |
| aaatccacct | ggctccggg | atccactgga | agtttatttc | tttagggttc | tatcccaacc | 178740 |
| agtcgcttaa | aaaccaagta | acacagacct | gaggggtggg | ggctggggac | tgcaactccc | 178800 |
| tctactcat | gggtggacagc | agtggggact | agggaggggc | aggagaggtg | gctgaagcaa | 178860 |
| ggcagcagta | atggggccac | gacgcacag | agccagctcc | gtcctctccc | agaccctggg | 178920 |
| gggagtcctc | gtggccttgg | gtggggagt | ggggacccc | cccaggccct | ccctctccct | 178980 |
| tcctcagaca | gcctccttcc | gggctcaacc | catttcttcc | ggcaggagac | taggccaac | 179040 |
| agagaggagg | aagtgggaga | ggaggacgag | ggaggggcag | gggtggcagca | caaatgaagg | 179100 |
| cagaggtgag | agggcgtggg | aaggccactc | cacccccaca | ccaccccag | agaggggcga | 179160 |
| gggaagccaca | ccatcacgca | gcattgtcgg | gggacaaggc | ggggtttaa | gctgaggggc | 179220 |
| ccggggcagg | cggggcctcg | ggcctcagtc | aaagccgtgc | cagtcgctgt | gctctgagtc | 179280 |
| gtattccagc | tcggcgccca | cacacttgac | accatccagc | agcatggggc | tgccgtgggtg | 179340 |
| ccggtctcga | agggcagggg | caagtgcagtc | ccccggcccc | acccgctggc | accatgcggg | 179400 |
| ccactaaagg | ggaccctccc | ccttcctca | gggatcagct | ggaggtaggg | acctgccaa | 179460 |
| gaggttgaga | acccctgagc | cgggcaagga | tcctctgttc | agccttggtt | ccctgaggag | 179520 |
| gacagaaaac | ctcgcagtcg | agcttctgca | tcctctctcc | aaccaggagc | ctcacgctct | 179580 |
| caccactgac | gcgggctctg | accctcacgc | gcacacaggt | gcccagtcca | ccttcgcagc | 179640 |
| ctagcagtat | ctcctcttgg | atggacgccc | ctttgtgcgc | cgaagtactca | cacttgacac | 179700 |
| tataacctgc | ccggggacag | gggcaattgg | tcagcacctc | ccccagcctc | cccgatcctg | 179760 |
| ccctcggcac | tcacggggccc | ctaccaccca | ccatgggggg | actgcattgct | aagccccccc | 179820 |
| aagaaaaggg | cagggatggc | ccttctggag | cccagagacc | catctcccct | agcaggggtg | 179880 |
| cacaggtctc | aaaaacccat | tcttcagtga | gctggacatg | cctccagcct | gtggggccacc | 179940 |
| ccaatgtggc | tccaagagtg | aatgaagtga | ggtccaattc | aggcttccaa | aagaaggtct | 180000 |
| ggctgttttc | tcccaaaaagg | aaggcaggga | gaggcggtga | tgaggagtga | ggggggcagg | 180060 |
| gcagggtagg | ctttgagcag | atccgatggc | aagaggtaaa | ggcctgaggg | gtgtcaatcc | 180120 |
| attgaagggc | aaagtcatga | agaggtcgtg | gacggggaca | tgaaactgaa | gctggagctc | 180180 |
| cacgaactgc | acaccctagc | ctctgccagg | ctgggagtg | ggcatggggc | agggcctgaa | 180240 |
| gtgagcctga | taggaacagc | agcctgaaac | ccaaggtctg | ggactgggtg | gtgcattagg | 180300 |
| ttatccacac | ggtgtgtgta | gctccagcag | gtttttctag | aggggttaggg | aggcaggagg | 180360 |
| gaaggctgga | ggcttcaaac | cagttcctca | gcagctccca | tcttgggttac | tgccccagg | 180420 |
| aggttaacct | cacacctagg | gcaggtacag | ggaagtatga | gtcctatggac | ttctgttacc | 180480 |
| gagcaggagc | aaggctctgga | tgccatgctg | ccacagtgc | ccacagtgc | agctcacaaa | 180540 |
| gaggtgccac | atccccgact | tgctgagctg | cccatccacc | ctagtggaa | gtaagggagc | 180600 |
| accocatgtt | ctcactccca | caccattcca | ggccctgctg | gaggagggga | cgcaccttca | 180660 |
| gggaacggga | ccacgctgag | gagcttgagg | tgacaggtgc | ggacaggtgc | ctcgcggaca | 180720 |
| tccttgctca | gcctgtgcac | tgggggcaga | gtgaaggtaa | tctcatacct | gtgcaggatc | 180780 |
| tctaggaagc | caacctacag | caggagaggt | gagggaggga | ggcaccttca | cttctctgct | 180840 |
| cccacaggcc | caccagctgc | tgccacctat | gatcccatgc | gccccagcac | ctgtctcccc | 180900 |
| agagagccca | gaaagcccaa | ctccctccat | cactccagct | gatccacccc | cagcccttcc | 180960 |
| acagagccct | tgagggtccc | agccctacag | cttgccacc | aaagtacctca | gcactgtctc | 181020 |
| cacgggcccc | aaagtcttca | gtggctcatt | tcatagacag | gaaaatggaa | accacaaaat | 181080 |
| gactgggcaa | aggtcatgtg | gtgtggccag | ggccaagtcc | tcctctagact | tcttggccat | 181140 |
| tctgtctatg | catggcaccc | atggagccac | caccactgc | ctacaaggga | aagttcaaa | 181200 |
| tcccatgggt | cattcacctg | tcccatcac | tgccagtatc | ccagggtcag | ggatgctggt | 181260 |
| cctccattcg | ctcataggat | gacaccaggc | cacagatagc | gtgggatgag | ggatgggatc | 181320 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| aaattagaga | tatgacagta | gaccagctgt | gtcagcccg | agctggccca | cagacctggc | 181380 |
| accagccact | gaactggggc | tccctggctgc | cctgctggca | ctggactggg | gatgcctgcc | 181440 |
| tgccaccact | aaggtctgtg | ccaatggccc | atactcacga | cccaagcaat | ttctgttgtt | 181500 |
| cattcacagt | gcttgggtgac | aaccaccctc | tgggacaaat | gcattctctg | gaaactccag | 181560 |
| tagtatgaag | ggtcacaagc | caggggtggt | gctgagcagg | ggggctggtg | gggggtgccac | 181620 |
| gccagcaggg | aaaggggtgcc | tgctgtactc | ccagcttgca | tgggcacaca | gagtcctgct | 181680 |
| cacagttact | ggggctgggc | agccccatcc | ctggggggcca | actgggactg | gctgcagaga | 181740 |
| gttttagcca | ttcaatggga | ccaggttgat | attgctacat | gacaaacttc | aatccatgtg | 181800 |
| ccccctcat | gctctgctgg | tagggggcac | atccctcaag | aaggtcttct | ggctgtacct | 181860 |
| gctaactgtg | accctgtgac | ccaggggtac | tacttataga | attttccagg | ggaaatagag | 181920 |
| atgtgccagaa | aaatactctt | cgctgccta | tttataacaa | tgaatacaaa | aacagccgta | 181980 |
| acaccagga | acagagtagc | agctaagtca | atgatgacac | atcatctgga | ccagtaggca | 182040 |
| gttgtaaaa | ctcgtggaaa | attattttact | gacacaggtg | ggagacaggt | cacaacagaa | 182100 |
| gcaagcaaaa | accagttgtg | acacggtgag | gccagctttt | gtttttggaa | acaaagcttt | 182160 |
| ttaaaaaag | ctggatccat | caccatagt | ttttcagaaa | aaacaaataa | taaaataata | 182220 |
| acattttaaa | aagctggaat | aagccaggtg | tggtggctca | cgctgtaat | cccagcactt | 182280 |
| tgggaggccg | acacggacgg | atcacgaggt | caggagatcg | agacctact | ggccaacaga | 182340 |
| gtgaaccccg | gtctctacta | aaatatataa | aaattagctg | agagtgtgtg | ccacgcctgt | 182400 |
| taatcccagc | tactctggag | gctgagggcag | gagaattgct | tgaacctggg | agctggaggt | 182460 |
| tgcaagtggc | tgagatggcg | ccaactgtact | ccagctgggt | gacagcaaga | ctccgtctca | 182520 |
| aaaagaaaaa | aaaaagctga | aataaacacac | atcagaatat | taacctgggt | tattttggga | 182580 |
| tagtatgact | gaccgtacgc | tgaagaggat | gcgtattgac | tgtgctttcc | tctcttgctg | 182640 |
| gtggatatga | aaactggaat | ggacattgct | atgacgggat | ggagctttct | ttcaaaagctg | 182700 |
| agcatgagtc | ttacataca | atctagcaat | ttcactccta | gatgactacc | caagtgaagt | 182760 |
| gaaaacttat | atccagggcc | gagcacagtg | gctcacgcct | gtaactccag | cactttggga | 182820 |
| ggccgacacg | gacggatcac | gaggtcagga | gatcgagacc | atcctggcca | acagagtgaa | 182880 |
| accccgctc | tactaaaaat | ataaaaaatt | agctgagagt | ggtggcacgc | gcctgtaatc | 182940 |
| ccagctactc | tggaagctga | ggcaggagaa | ttgcttgaac | ctgggagctg | gaggttgacg | 183000 |
| tgagctgaga | tgccgccact | gtactccagc | ctgggtgacag | caaagactccg | tctcaaaaag | 183060 |
| aaaaaaaaaa | gctgaaataa | cacacatcag | aatattaacc | atgggttattt | tgggatagta | 183120 |
| tgactgtggg | tactttttaaa | tttcttctca | ttttctctgt | cttccaaatt | ttctcatgtc | 183180 |
| tattaaagaca | tgacttggatt | ttgcaattag | gggagagagc | tatttttaag | tgtttgtttg | 183240 |
| ttcactttac | tctcttgggg | aaagctgtcag | ctgtaggaca | aaagagtggga | actcagaggt | 183300 |
| tggaacagaa | tggatacaaa | cctggggccc | ccaggttctt | ttttgtttgt | ttgtttgaga | 183360 |
| cagagtcctt | ctgttgccca | ggctggagtg | cagtgccaca | atctcagctc | actgcaacct | 183420 |
| cagcctcctg | ggttcaagca | attctctctg | ctcagcctcc | caagtgtctg | ggactacagg | 183480 |
| cgctgcccac | cacgctcagc | taacttttgt | attttttagta | gagacgggtt | tcgacctatt | 183540 |
| ggccaggcta | aacctggggc | ccttttgagc | agcgcagcag | acccccact | cagggccatc | 183600 |
| tcaatggggc | agacctcccc | cgacccaagg | gcactcctgt | tcaactctgg | acccttgatt | 183660 |
| tatttgaccaa | cctgaaggcc | tcagcttcca | ttttccccaa | tcagctcca | ccaagcagaa | 183720 |
| aacagggggg | gataatacta | cttcaaagag | cctaactgag | ccagagattt | ggcaaaagag | 183780 |
| ggttaaaaa | aaagttgac | cttgttatgt | ccatagttag | cttcacacct | gtatcacata | 183840 |
| cataacttct | tccattctca | caccacaact | aggggatgtg | atttgtctcca | cttgagagaa | 183900 |
| aagaaactca | ggtagacctc | ccaaggtcac | agagccagaa | tggtgtggct | agacctgaac | 183960 |
| tcaggccaact | ggcacacac | ccacgacact | gccttccatt | gctattgtct | gggtgaagca | 184020 |
| gatgacagca | gtggctctgc | tcactctgat | ctggatgaca | ttctaaatgc | ttctcatctca | 184080 |
| ttgaaactcc | actgcctccc | atctgacaga | tgggggcaac | aaggcacagg | gaaaggcaca | 184140 |
| gacatgccta | taagcccggc | tgagaggaac | ggggataggc | acccagaagc | caggcagagc | 184200 |
| cggggagggg | aagaagctgc | tgatggggta | ggtgtgcatg | taacctgacc | agaaagctgc | 184260 |
| tgtaactctc | ctgggtgacc | atgacacccg | agtcattgac | cttctcatca | aagttggacgt | 184320 |
| ggctgtggga | tcctctcgca | tgtggccctg | ccgcaaaagc | gataactccg | actctgggct | 184380 |
| tgttgcctag | gaagagtggg | agagggctct | gagggcttct | cctgtcccca | ggaaactcct | 184440 |
| cccccttgct | cctcgtcac | ccccaaagct | gcccttgaca | tcactccagct | cccactgggt | 184500 |
| gggtctcttt | cagggctaga | tgggacactag | gatcatgagg | cttgaggcct | cccccgccgc | 184560 |
| cccgacactgc | ccctcccaca | aaaatccatc | ccctgagagt | gagcgtgggg | ggactcaggg | 184620 |
| actggctaat | atgacccctg | cttggagggg | gtggggctgg | gtccagagcc | aggaagggac | 184680 |
| agtcttccag | actcccacca | agccagggga | gcgtgggact | cagcccagtc | ctctggatag | 184740 |
| cctgcccaagt | gctctccttc | acagtgcaaa | gctcccatc | cctggggcct | accctaagc | 184800 |

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|------------|--------|
| tgtgtcagtg | catgagggccc | ctgcctgccc | atcctcatga | cccaaacctg | aaaggcgagg | 184860 |
| gacaggaaca | ggctctgggg | gatctgcctg | gggtggcaga | aacagggggg | atcaaaaaac | 184920 |
| acactcaggg | cacctggggc | aggctgcagc | tgccactaaa | cctcacttgc | ccttgggcag | 184980 |
| aagaagagag | aatccaagga | agcttccaga | cccactccaa | gaccccttcc | cttctcttgt | 185040 |
| ctagcccacc | atcaaggccc | tcagctgtct | ctgagctggg | tcgggttctc | caaggccaga | 185100 |
| aaggggcagc | ggatgagaca | ggcagggtag | gcgggagccg | agctgaaggc | ctggactgga | 185160 |
| tgaatgctca | ggctgtatag | gccgagagca | cgcccttagt | ctctctggac | cctctggccc | 185220 |
| tagggcctcc | caccaagagc | cctgcccagg | gtcctcttgg | cgggaggagc | tgatccagct | 185280 |
| ttgagccggc | ctcttgata | ctcggaggcc | caggggacca | gcctggcaca | tcacaatcc | 185340 |
| ctggtagcc | ccacagaacc | ctcaaatctc | accagatcac | cctggcctaa | aaccctccct | 185400 |
| tatccctctc | ggactctgaa | agaaaaccaa | actcccttcc | acagcctgct | gtggccctcc | 185460 |
| cacatcacca | gtctcagtc | tcaccataag | ccctcggct | cccccaccc | ccacagtga | 185520 |
| acggcagggg | cacaagagca | cacacctcca | cagaccagcc | taaggagtca | gcagggtctc | 185580 |
| gcaggggtcc | tcttagggct | tgctgtgagc | ccagaacctg | gcacaaacac | tgacatagtc | 185640 |
| ccaatcacat | tcgcacacca | gcacacacgc | actcacacat | gcacacacac | acacaggtgg | 185700 |
| tgccggccct | acctgcagca | atggccctcc | taccagggcc | caccggaaga | gcttgtgcaa | 185760 |
| gtctctacc | gccaggaagg | cacttaccct | tggtggctgc | agccatggac | gcctccctgc | 185820 |
| ccacgcagct | cctgccagac | accgccactc | acgctcagca | gcctccatgc | ctccagggac | 185880 |
| accagcggga | gcctgaaggga | tagggagtg | ggaggcgagg | ggtcagggcc | accatggac | 185940 |
| ccatggctca | gggacacac | aggagctccc | tttaggaagg | actcaaaccc | ctaccgctct | 186000 |
| caggcagcca | agtggtgcct | gtactctcta | gcacgggggc | ttctgcccac | cctctaccga | 186060 |
| ctgcccacat | tcaccaagag | ccctcacccc | ttcctgatac | ccagtgact | cagggtcctg | 186120 |
| ccagcatgtg | ctgagctgca | gcttctaggt | gccaaagcaa | cagcatagga | ctcctatccc | 186180 |
| cagcaccccc | agagactggc | aggaggggca | gcctggaaag | gaggacttta | ttggatattt | 186240 |
| ccagggagcca | gttctgtgtg | tggtgacccc | aaaggctggg | cctggagcta | ccttatccag | 186300 |
| gtcacacaag | caatgcagct | ggctgcggtc | agaggcaaca | gggtgtctga | ttggtggcaa | 186360 |
| gaacaggggg | ccacataaag | taagcgtgcc | cttagagctt | tcctcctg | tgatcgttca | 186420 |
| gggcatatg | caaaaacagat | gcctgtggag | ggggtgtgcc | cccactgaa | cccattcag | 186480 |
| acctgcctc | gagtcctagg | ctcgcctccg | cctcagctcc | caacaggcag | cctcagaca | 186540 |
| ggggcagtga | gaggcacccc | aggaagcctc | caatggggcg | agctgggagc | atctgagcat | 186600 |
| caaaaagaat | aatgaagtca | actgatggaa | acagatggaa | acataaaag | cagtgctatt | 186660 |
| acagagatta | ttagaaaaga | agagagaaag | caaaaacaaa | aaacccaagc | cctcaaaagc | 186720 |
| cctcaattgt | caccactgac | ggttagcagtg | ccctctttac | ctctgaaagg | gaggattaaa | 186780 |
| gaggaagaat | tcagcctgtc | tcttgtaacc | aaattgccct | ggtggttgat | tggtggtgat | 186840 |
| ggaaagctct | tcttccggga | agaattcttg | ctaataaata | acaaagaaat | gaccaaatgc | 186900 |
| taagtcatte | tgcaaccctc | aaaggaaaca | atgctggagg | caacaagaac | cagtggtatg | 186960 |
| taaaccagag | gggaagggtg | acaggaagca | gatactocaa | ggcgcccaag | cgtcactgac | 187020 |
| agaggacctg | gtatggtttt | tggttttttt | ttttttctga | gacggagttt | cactcttgct | 187080 |
| accagcgctg | ggatgcaatg | gcacgatctc | ggctaactgc | aacatctgcc | tcctggggtc | 187140 |
| gagcgactct | cccgccctag | catcccgagt | agctgggact | acaggcgccc | gccaccacgc | 187200 |
| ctagctaat | tttgtatttt | tagtagagac | agggtttcac | catgtttggc | aggctgggtc | 187260 |
| cgaactcctg | acctcaagca | atctgctcac | ttcggcctcc | caaaagtgtg | ggattacagg | 187320 |
| tgtgagccac | catgcctggc | cgaggacctg | gtgtttaaga | ggaatggcac | aagtttgcaa | 187380 |
| tggaggtatg | tgacttctcc | cagtcacaca | ctggctctat | ctctctactc | ctgcaggaca | 187440 |
| gcctgtcacg | ctgatgcctc | ctgaagttag | gcagcccaaa | gtcacagcat | cactgacaag | 187500 |
| ggattcctgc | caaaaagggt | taacctggaa | tctaaaccag | cctctaactt | caactctcca | 187560 |
| tttgctaaaa | acacagagga | gaggggaca | aatttaatga | ctccatgaga | aagcaatgag | 187620 |
| acagagccag | aaggcaggat | attctacagg | atgactgacc | tgctttttgt | ttttgttttt | 187680 |
| gttttcaaat | gtcaagaaga | aaaagaaagc | aggctgggaa | cggctggtca | cgctgtaat | 187740 |
| ccagcactt | tcacggatca | caaggtaagg | agttcaagac | cagctctggc | aacatagtga | 187800 |
| aaccagctct | ctacccaaag | tacaaaagtt | agccgggtgt | ggctggcggc | agctgtagtc | 187860 |
| ccagctagtt | gggaggttga | ggcaggagaa | tcacttgaat | ccagaaggcg | gaggttgacg | 187920 |
| tgagccaaga | tcgcccacgc | gcacttcagc | ctgggtgaca | cagcgagact | ctgtctcaag | 187980 |
| aaaaaaaaa | aagaagaag | aaaaataaaa | cgaggagat | tattctagat | aggaaagcag | 188040 |
| caaatcacag | ctgattacac | agctgatggg | aaggttgctg | cctgttgtgt | tttggttttt | 188100 |
| gtttttgaga | caggtgtctc | ctctgtccag | cccagctg | agtcagctc | actgcacccg | 188160 |
| taatctccca | ggctcaagca | atcctcccac | cccagcctgc | caagttagctg | agaccacagc | 188220 |
| cacgcaccac | catgccccgc | taatttttgt | agagacagag | ttgccacagg | tggtctcaaa | 188280 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ctccggagct | taagcgatcc | accacacctca | gcctcccaaa | gtgctgggat | tacagacgtg | 188340 |
| agccgccagg | cctggcaaat | aaattttttgc | aaataaaagt | ttactaatag | aacactacac | 188400 |
| ttcatctact | tacatgttgt | ctgttttttgc | actaaaacaa | agttacttcc | aagtagttat | 188460 |
| gatagagacc | atatggcctg | caaagctctaa | accatttatt | atttggccct | tcagagcaaa | 188520 |
| agtttgccaa | cccttgttct | agatcaacac | acacttaaca | gacaaaaaaa | ggataatctt | 188580 |
| caggacaact | gacctgtttt | tttcaccaca | tcagctgcaa | gaagaaaaat | aaaggagagg | 188640 |
| gtgtttcgaa | ctacctatca | aagaaagcat | tttggggggt | gccggcgctc | gcggctcaca | 188700 |
| cctgtaatcc | cagcactttg | ggaagccaa | gcgggtagat | caoctgaggt | caggagtttg | 188760 |
| ggaccagcct | ggccaacatg | gcgaaccccc | atctctacta | aaaatacaaa | aatcagctag | 188820 |
| gcatgggtgt | gtgcacctgt | aatccccagct | actcaggagg | ctgaggcagg | agaatcgctt | 188880 |
| gaacctggga | ggcggagggt | gcagtgagcc | aagatccgcg | cactgcactc | cagcctgggt | 188940 |
| gcagagcgga | gactctgtct | caaaaaaaaa | agggcgggtg | aggtggctca | | 189000 |
| cgctgttaat | cacagcactg | tgggaggctg | agggcgggcg | atcacctgag | ctcgggaggt | 189060 |
| caagaccagg | ctggccaaca | tgggagaaacc | ctgtctctac | taaaaaatc | aaaattagcc | 189120 |
| aggcgtgtgt | gcgcattgct | gttaotccgc | tactcgggag | gctgaggcag | ggaatctgct | 189180 |
| tgaacccagg | aggcggagggt | tgcgggtgagc | cgagattgct | ccactgcact | ctagcctggg | 189240 |
| caaaaggagg | aaaactccgt | ctcaaaaaaa | aaaaaaaata | agacatttgg | gggattctgga | 189300 |
| aaatgttaaat | ggcctggata | ctagagtaaca | ctaagaaatg | agttgatcat | ggaattgtat | 189360 |
| tatgtaaaga | attatgtcct | tatttttgtca | gggaaacata | ctgaacaact | tagagatgag | 189420 |
| tgacactggc | tcatcagtc | agggacagaa | gaggaactac | aggggtggcct | catagtttag | 189480 |
| ggctgggcct | gtggaggcca | tgatctcaag | gtccctctgg | tgcccaagctg | ctctgggagc | 189540 |
| agatgtggcc | tcacctcgct | ccccacttct | caatgtgggc | aaagtccacc | caggcccaag | 189600 |
| ccctgcctca | tgggcagggt | gttcacagtt | tgggtaccac | caggagcccc | ctttggccaa | 189660 |
| tggacacagg | cagcaaacccc | ctcttggagg | tgagcttcca | ggccccagcc | cagggtgcag | 189720 |
| gaagtggagc | tacacccatt | cacctctggc | cagggccact | gtggggctcag | ttgcctcagc | 189780 |
| cctggaaagc | tcagcttgct | cccagggaac | ttggtcttgg | agagcagcct | gccgtaaccg | 189840 |
| ccccccgact | acagctgccc | ccaaagctgt | gaactcagag | tatatgacaa | atggccaagc | 189900 |
| acaagagtgt | gaggccctcc | tatcctcagc | aggtgctccc | agccacacct | ctgcagccag | 189960 |
| gaggcagtg | cagcccgctg | tcctcaaggt | cccccatct | cccaggatgc | ttcctgaagg | 190020 |
| ctcaggttaa | actgcaccag | tggcttggtga | gcaaaaggctg | caggcccttt | tcacaaaaaca | 190080 |
| ctctgaactg | gtccatgagc | ggcacccctc | ctgtcttctc | cccaacctac | tgccctgacc | 190140 |
| tcgccccgct | ctgcctttgc | ccaggtcabc | cttccagcca | ggaacaaatc | gcctccctcc | 190200 |
| agccacacct | tcgacagggt | ctcagcccca | acatttcttt | tttaaggatg | cacaccctgc | 190260 |
| aaatcccagg | acaggactac | agctcttttt | ggaatccctt | ttctcatgtc | tgggatcatt | 190320 |
| tgatatccct | cttctccacc | tlacaagggt | agtgactgtg | tttttgtcaa | acctgtgtgc | 190380 |
| cccggcactc | agcgcagcac | cagacaggga | ggagctgtgt | ttgggttatg | tttgttgaat | 190440 |
| gaatgaccac | atcacatttg | cttggagggc | ctggccaagg | cccgactctc | agccctaaat | 190500 |
| gatattggtc | gggcccagtg | ctgcactggg | ggctctgagg | aagcagaact | ccccacccca | 190560 |
| ctcttttggc | tcacccctgc | ccctgggggg | acttccctct | ccctggcccc | caccaagtgt | 190620 |
| cctcccactg | ggactcctca | ggccttctgc | cagtcagctg | gttacctgtc | catgcttgct | 190680 |
| ttgagaggga | catgccccat | gcccactgag | ggtaggaccc | acatctcatg | cagtgcaggt | 190740 |
| actcagtaaa | gggcaaacac | tgaggggcctg | taacctctgt | gatagtgaca | acatagaggc | 190800 |
| aggaagcaag | ggacttcagg | aacctgggaa | aaaccaacct | ctcttctcaa | ctgttctcaa | 190860 |
| tggagaacct | gctgggtcgt | gtccctgggg | agctagactc | tgtgtcacac | acaattctcc | 190920 |
| atcagaaaca | ggcaaaagg | gcactggggc | tcctcgtgct | cagctgtgtc | ccagagaaac | 190980 |
| agctaggcca | ccacacaacc | ccaagccttg | ccctgcctcc | agcctctccc | tgtgccaacc | 191040 |
| aaaggcctcac | catggctcaa | tcacatgggc | ccccagaggc | agcagctcgt | gttcaaatcc | 191100 |
| aggatctgcc | cttgtatgct | ctgtaggtga | gttctttcac | ttctctgtcc | ctgtttctca | 191160 |
| atgtgtaaca | caggaataag | cagtggtcct | ttccttccag | ggtttgtctg | aactgtgctc | 191220 |
| gaagctctgt | gacaatgcac | cccccatctc | tgcaaaagca | aaoccccaaa | ggcctggttt | 191280 |
| ctaaagcaat | agcagtttca | gcagacccgt | cagagaggca | cttcaggcca | atcctggagg | 191340 |
| agccaggagt | gacctttaga | gtggggccca | gggacgtcag | ccttcttgga | aaacagctca | 191400 |
| aggggtgagg | gggcctccct | cctcctgcct | cccccttctc | ccactcccaa | agcagccagg | 191460 |
| tccttagggg | gggtccagaga | acagatgtgt | gtccccctaa | ctcagggggg | ccagaggggg | 191520 |
| tcacaaggaa | gattgtcaga | atgaacatcc | tgggaaactg | ggaaatgact | agggaggaac | 191580 |
| atggtgcctc | cccctcgaga | aaaaaaaata | tacccttccc | catgagatgg | agtgctcagca | 191640 |
| agcttccagg | ccccagccca | gggtgcagga | agaggagcta | caccattcca | cctctgggtc | 191700 |
| gcatgtcccc | aacgtctgag | acctcatctc | tcatctcttc | ttcagtcctc | attctctattg | 191760 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| tggtttcagg | tctttctctc | tgagccagga | actccgcaac | cttcaccctt | ccctacctct | 191820 |
| tcctctccta | cccagcctg | gggtcagtc | ctggctcaag | cactcagtc | agcagaagca | 191880 |
| ctgtgtagcc | tccatttaaa | gctcagcgt | gtgaaaagaa | cacccattga | ggccttgaga | 191940 |
| tggggccaca | ctgaccgcct | gactctcagg | actggacaca | gcagaggcca | cacatactca | 192000 |
| gaacaaagcc | tggaaggcca | aggctggagg | tcagtagttg | tggcagcttc | acatacactc | 192060 |
| agctttaatg | tgatttaatt | tctctctccc | tccagtgggc | caaagggtga | aagataagta | 192120 |
| tggctgttct | ctctcctctc | aacagtgagg | tgctgggggt | gggggtgggg | gaatatggag | 192180 |
| aaggggacct | caccaccacc | acctctctgc | ctccccaaac | agtgtctgcc | tctctctgcc | 192240 |
| agcattctcc | ccactttgcc | ctcagctagt | gggtgcttag | ctccagata | gcatacccca | 192300 |
| cctaggccct | gccttggggc | tgtagtccag | agggtcccaag | aagcagaggc | caggctggat | 192360 |
| ccaggggggt | agccaagggt | aggggtgggg | cacacaggat | tatctccag | ggacagggtg | 192420 |
| gctgcctcgt | agctcaggat | ggatagaatg | tgggggggata | tccagctaca | ttttccctcc | 192480 |
| acaaaagacc | agaatgggag | ggggatgggg | tgtctccccg | actttcttca | actccccgga | 192540 |
| gcagaaaaat | gcctctacct | cactttccag | tgccaagatt | caagaagaaa | ggcaagcgga | 192600 |
| gacttccctt | tctcagctcc | gtctactaa | tggaacaacg | ggtccagaa | ctaaatccag | 192660 |
| tccctctctc | ttcataccac | cgggagggag | gtgcagccca | agcccccgag | gccccaaagg | 192720 |
| tccagggtga | ggacccttta | tctctctccg | cagccatccc | tgtgggtgtg | gcacccccgc | 192780 |
| cacaccccat | tcttgtcttc | tccaggggag | gggggaaatg | taactggaca | tcccccccat | 192840 |
| ccaatccatc | ctgagctcgc | agggcgccgc | tctgtccctc | ggagataatc | ctgtgcactc | 192900 |
| ccaaccttca | ctcactcggg | ctgacgcagg | agtctccgga | gcccgcactc | ccagacatca | 192960 |
| ctgcccctct | tcttgagggt | gctgagggtc | ggaggctcag | agatgtctact | ggtccaagggt | 193020 |
| catgcagcga | ggcagcgcca | atgaacaggg | tctgtggggag | gagggggcgc | tgaccccat | 193080 |
| acgccccccg | cctcactacc | gcactcatgc | ccccggacaa | tgcgtctcgc | gaaaaacacc | 193140 |
| cagctgccac | cagttaacgg | tcactgcgcc | ccgggaacct | gacatcactc | tagttcagggt | 193200 |
| gctgggggag | tctccaggcc | cggcccccaa | cagcggagcc | ccccaccctc | agccctgcca | 193260 |
| tccagggccg | tggggtccca | ggcactgact | gctcaggaca | gggcggggag | cggggagggg | 193320 |
| gacctcggcg | aacgggaaag | aaccgggagg | cagggagcag | aggggtcgcc | tcaccttgcg | 193380 |
| gggtgcaccc | cggggccggg | gagggcaggg | acgaccactg | cagcggcgcc | ggctgcagga | 193440 |
| gctcaacggc | gagcacagag | aagggagccc | cgcgcgcggg | cgccctctcc | ctcggcacgc | 193500 |
| ccccgcctcc | gccattggat | tgtatctggga | gggtggggcg | aggggaacct | cggacccaatg | 193560 |
| agcggtctcc | aaagaacggc | caacttcgca | gggcgcgcta | cgtaacgtgc | cagggctcgcc | 193620 |
| gaggcagcgc | cctgctagtc | cgcgcctgcc | gggcgagctc | tccggaaggaa | cagcggcagg | 193680 |
| cgcgcccaact | agggcagggg | ccagaaccga | ccactcgaag | aggggaaggag | agggcctcgt | 193740 |
| ataggccccc | ccccgcctcc | tctctccgcc | tgggggtag | gcctctagc | cttgaacctt | 193800 |
| gcttaggaacg | cacctccctt | gggcccctcg | ctctcgggag | ggctgtcggg | cgcgtctcgg | 193860 |
| ggctgggtgg | agctcccgaa | gggtggcctt | ctccctgggg | ttccacgcgc | gcttcggcca | 193920 |
| tcgatacggg | ccgtgttggg | ctcgtttcagg | agctgaggaa | ccctccatca | ctcctgtttc | 193980 |
| gacccccagg | tttggaacct | ttccccctcc | acccccatcc | ctgtcttgaa | agaagcaacc | 194040 |
| cccgctcggg | ccgagacgc | gtcccggggt | cctggccggg | cctggagaa | catcagaaca | 194100 |
| aagaaggcac | gcgggctggg | ggctgggaga | gcctgtgacg | cggccccggg | gaccgcagcc | 194160 |
| tctgtctccc | gtctccatgg | aggcggtcgc | catggcagcg | agatgcgcct | cgctcagcac | 194220 |
| cgcgggggtg | gatgtgggcg | cctgcaatga | gccagggagc | gagaggcgtg | gcctccgggt | 194280 |
| ctcggggggg | tcttgccggg | gctctccggc | cgcgggttcg | cgaacacccc | acctatactt | 194340 |
| cgcccggtgg | gacgggattc | ccaaagtgc | ctcagtaagc | cgtccggagc | acgcagcgcc | 194400 |
| ttgcttccaa | cggaaactaga | gggcggccga | agggcagcct | agggcagcct | cccttcagca | 194460 |
| gggcgcgggt | cgtctgcctta | aaggagccct | caagtctgcc | acccgtgtgc | ccataacctg | 194520 |
| tctgtctgat | tccagttctg | acactgttgg | caaatttaact | ttctctgagct | cttgtttttc | 194580 |
| tcgcgtctct | ctctctgctc | aaagccctct | gggactgcct | ccagtagcgc | ttcacaaact | 194640 |
| tcagcagcac | tttgggtgac | tcatgtgcc | tccgctttga | gagacagcgc | ttacattgag | 194700 |
| agcttttcc | attctgatct | cagcccatcc | actcttgccc | actctaccca | ctccccacc | 194760 |
| actacctttc | gaccttgtct | tgaagtctta | aagttggcag | acggttgggc | caaagtggcc | 194820 |
| cgtagggtga | ttttgtttag | ctcacatggg | gctttttaaa | agaagtccct | ccccctcctg | 194880 |
| ctctctcgc | ctcttctctg | aatcaggaat | ctggtttcgt | ggcttttgaa | atatcagaag | 194940 |
| atctgacaac | gtgggctcac | tttccagcct | ggcattacaa | gcccccttaa | gagcacataa | 195000 |
| atttaccata | gtccccccta | ctccctgttt | ctcacgggtta | gttacctctc | gggtctgtgc | 195060 |
| agacatttgt | gtccccctcc | acctcatat | ccccacttcc | aaccttggag | gcctccccgc | 195120 |
| ctctctctc | atagccaaaa | ttcaagccct | cttcaaccaa | gataagtttt | gcctatttat | 195180 |
| ttattttatt | atttttgaga | cagagtctca | ctctgtcgcc | caggcttgag | ttcagtggtg | 195240 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ccatctcggc | tgactacagc | ctccaccccc | tcccccgccg | ggttcaagcc | attctcctgc | 195300 |
| ctcagcctcc | caagtagctg | ggattctagg | tgcgcgccac | cacacccagg | gctaattttt | 195360 |
| gtatttttag | tgggacagcg | gtttccaccat | gttggctcagg | tgggtctcaa | actcctagcc | 195420 |
| tcaagtgatc | tgctctgcctc | agcctcccaa | agtgcctggga | ttacaggcat | gagccaccgc | 195480 |
| gcctggccca | agcctgcctt | tttcaaaagt | ctttcttagcc | aaactctcaa | aaccatctct | 195540 |
| gggtgsggct | ctctcaaaac | caaactggct | gctaactgca | cagcccgggg | ttattttccga | 195600 |
| atgtgtcaact | cccttcagca | aacattttact | gaatgtatct | gtcaggttaag | ttattgtgtct | 195660 |
| aggctctgtg | gagtggaagg | cactcacagt | ttagtgtagg | agggacacca | cacacaaaaga | 195720 |
| actgtggctc | ctgtccttg | ccaaaagtct | ctcacactcc | caggcatcgg | gcagtctggc | 195780 |
| ccactgggag | gagctgaat | acagagctgc | ccatgcctag | ggatactaca | tgctctcagg | 195840 |
| gaagggccaa | gatggtttctg | gaaagcctcc | aggagaaaat | tagactcttt | gtgtcaggag | 195900 |
| gaagtacagg | caggagttac | ctggctgccc | catcactgcc | caatgtgtct | gtgttaccac | 195960 |
| acgagtagtg | cccaggtct | ggggcttct | gtggcgaggg | ccttaccaaa | gatagatggc | 196020 |
| tctgcaaggg | aaacacacag | cccctaccag | gccccctgga | gacagaggca | ggccaagaag | 196080 |
| aatcaggaac | aactccaggt | taggagaaca | gatggagcac | gtataaaaac | tccaactcac | 196140 |
| gcccggcgctg | gtggctcacg | cctgtaatcc | cagcactttg | ggaggctgag | atggggcggt | 196200 |
| catgagggtca | ggagatcaag | accatccttg | ctaacacggg | gaaaccccg | ctctactaaa | 196260 |
| aatacaaaaa | attagccggg | catggtggct | gatgcctgta | atcccaacta | ctccgggaag | 196320 |
| tgagggcagg | gaatggcgtg | aaacccggga | ggcggagctt | gcagtgcagc | cgagattgcg | 196380 |
| ccactgcact | tcagctttgg | gcgactgagc | cagactccat | ctccaaaaaa | aaaaaaaaaa | 196440 |
| aaatccactc | accatgagaa | ctcgtgtgact | tttaaaaaaa | aacaaaaaaa | cacaaacact | 196500 |
| gctacaactc | caaacacacg | gtttttttgt | ttttttttta | tttttcagag | acaggggtctc | 196560 |
| actatgttgc | ccaggtctgt | ctcaaaactc | tggcctcaag | cgaatccccc | ctcttgggct | 196620 |
| cctaattgtc | tgggaccaga | ccagaggcac | aagccactgt | gcccagccct | ctattctttt | 196680 |
| taattaaaaa | ttaaacaaat | ataacccat | tggaaaaactg | cttggcagta | cctaataaaa | 196740 |
| atgataaatg | aaatgagtaa | ttccactttt | tttttttttt | tgagacagag | ttctaactctg | 196800 |
| tcacccaggc | tgggaatgcag | tgggtccgac | tcagctctct | acaacctctg | cctcctgggc | 196860 |
| tcaagcactc | ctcccacctc | agcctcccca | gtagccaggga | ctacaggcac | atgccaccac | 196920 |
| accagactaa | tctctgtatt | tttttttttt | tttttgagatt | acaggcaatg | ggcaaccacgc | 196980 |
| cggctaaatt | tgtattttta | gtagagacgg | gtttctccat | gttggtcaag | ctagtctcga | 197040 |
| actcccaaac | tcaggtaaac | tgcctgcctc | agcctcccaa | agtgctggga | ttacaggcat | 197100 |
| gagccaccgt | gcctggccaa | ttttttttgt | tttttaatat | agacgggggt | tcaccatggt | 197160 |
| ggccaggctg | gtctcgaaat | cctgacactc | ggtagacgac | ccactctaga | ctcccaaggt | 197220 |
| gctgggatta | caagcttgag | ccaccatgcc | cagccatggc | atgtgtttta | ttttaaaaga | 197280 |
| aactggcaaa | ttgtttttca | aaatgactgt | acttaactac | actccatta | gcagagtata | 197340 |
| agagtttctg | ttgcttcaca | tcctcatcaa | cataaaaaat | tgcgaatttg | taaaaaacgc | 197400 |
| tacaccatgc | taacaaaaca | aagcagggga | acggactacc | agtcttcagc | ctctcctctg | 197460 |
| agccatgcgc | ttgtagaggga | caatttcaca | acagctatca | agattacaaa | tgacacacaa | 197520 |
| aaaaaagaga | aaacacacac | acacacacac | acacacacac | acacagccaa | catgggtgaa | 197580 |
| cccatctctc | actaaaaata | caaaaaatta | gctgggcgtg | gtggcagggtg | ctctgaatgc | 197640 |
| cagctaccaca | ggaggttgag | gcaggagaat | cgtctgaacc | cggggaggcag | aggatgcagt | 197700 |
| gaggcaagat | agtgctatgc | cactccagcc | tgggcccagc | cgcgagactc | tgactcaaaa | 197760 |
| aagaaaaaag | agagaggaaa | aaaaacacaa | caaagattac | agatgcatat | attctttgat | 197820 |
| ctaggaaccc | cagttctcga | atttatcatt | tagatatact | tcoccatgat | actaaaatcat | 197880 |
| gagtgataga | ggttatttat | tgttgcatto | ttgtgatagc | aaaagatttg | aaataactca | 197940 |
| aatgtccata | aatagaagac | ttcttaataa | aatttggtac | agacaatata | atgttatatg | 198000 |
| gctttaaaaa | gaaaaaaaaa | aaagggaag | ttttcaatgt | cctcccaagt | agccaggatt | 198060 |
| acaggcgctga | gccaccatgc | cgggataatt | tttgtatttt | tagtagagat | gggggttttg | 198120 |
| catgttatac | aagctggctc | cgaactcctg | gcctcaagcg | atccaccacc | ctcagcctcc | 198180 |
| caaaatgcta | ggattacagg | catgagctac | cgtgcccgag | ctgcataagg | aagatttttg | 198240 |
| attagctagg | attctccaga | gaaacagaa | taggaaagag | gcagatagag | aaagactgac | 198300 |
| atattataag | gaattgactc | acacaaatat | ggagctcgac | cagtcoccaa | agctgcagg | 198360 |
| tgggttgagc | agttggagac | ccaggagagc | caatggttta | gttcaggtct | gagtcoccat | 198420 |
| gcctgagagc | caggaaagtc | aatgatgtat | ctccagttca | aagggcgagc | ggtttgagac | 198480 |
| caagggaagaa | tcaatgttgc | agtttgatcc | caaaggcgag | aaaaaagctg | atgttccagt | 198540 |
| tcaaaagacca | ccaggcaata | aaaattctct | cttatttggg | gagggccagc | ctttttcttc | 198600 |
| tatttaggtc | tcaactgact | ggatgtggcc | cactcacatt | agagagggaa | atctgtttca | 198660 |
| ctcatctgct | gatttctttc | ttttttctct | ttgtttttga | aatggattct | cactctgttg | 198720 |

| | | | | | | |
|-------------|-------------|-------------|--------------|-------------|--------------|--------|
| cccaggctgg | agtgcogtgg | catgatcttg | gttccectgca | acctccacct | cccgggttca | 198780 |
| agcgattctc | ctgcctcagc | ttccctgagta | gctgggggtta | caggccaatca | ccaccacgccc | 198840 |
| cggtcaatttt | ttgtattttt | agtagagacg | gggttttcgcc | atgtgtgacca | ggctggctctt | 198900 |
| gaatgcctga | cctcaggtga | ttagcccgcc | tcggcctccc | aaagtgcctgg | gattcacaggc | 198960 |
| gtgagccacc | gtgcccggcc | tcactctgctg | attttaaattgt | taaccttagc | cgggtgcgggt | 199020 |
| gggtttacccc | tggttttccca | gcacctttggg | agggccgaggc | ggggcggtcca | ctcgagctca | 199080 |
| gggtgtcggg | tcaacatggg | gaaacctgcc | tctactaaat | acataaaaaa | tagctgggca | 199140 |
| tggtggcatg | agcctgtagt | cccagctact | caggagggccg | aggcacgaga | atacactgaa | 199200 |
| cccaggagggt | ggaggggtgca | gttaactgaa | attgtgccac | tacactccag | ccagggtgac | 199260 |
| aaaacaagac | tcttctctca | aaaataagag | taagacaaat | gtgcatgttt | gtttatagtc | 199320 |
| acataaagtc | attccgggaag | gatacataaa | aactaaagat | agaggttccc | tatttagagat | 199380 |
| tgctcgggaa | taagaagact | tttactcttt | tatgtgtttt | ttttaaatca | tatatattgtt | 199440 |
| ttatctattc | aaaacattaa | gtaacaaaagt | ttaaataccc | ttgcatgcct | tcctctcctg | 199500 |
| taggtgttggt | aacactggag | ggaaacaaatc | atgttttata | ctcacttttg | tattctcgtgt | 199560 |
| atttgataca | tagtaggtac | tcaataaata | gaggaaagaa | cgaatgaatg | aaccagaaga | 199620 |
| tagcagattt | agtgtcgggt | agagagagga | gagggggccag | gcacagtggc | tcacacatgt | 199680 |
| aatcccagca | ttttggggagg | ctgagggcaag | agggctcactt | gaacctagga | gttggagacc | 199740 |
| agctcgggta | acaaaaggaa | atcctgtctc | tcacaaaaaa | aaaaaaaaaa | aaaaaaaaaa | 199800 |
| aaaagcccag | gagcagtggt | tcacacctgt | aatcccagca | ctttggggagg | ccgaggtggg | 199860 |
| gtggatcacc | tgaggttagc | agttcgagac | caactaggcc | aacatgggtga | aacaccatct | 199920 |
| ctactgaaga | tacaaaaatt | agttggcaggt | gctgtgaatt | gctgtgaatt | ccagcaactc | 199980 |
| aggaggctga | ggcagggagaa | tagcttgaac | ctgggaggca | gaggtgtgcag | tgagccgaga | 200040 |
| tcactccgct | gcactccagc | ctggagtaca | gagtgagaca | tctcataaaa | aaaaaaaatt | 200100 |
| agctggatat | gggtggatatg | acctgtcctc | ccacctattc | tagaggctga | gggtgtaaaga | 200160 |
| ttgcttgaac | ttggggaggt | gggggttcgag | tgagccaaga | ttgtggccat | gcaactccag | 200220 |
| ctggggcgaca | gagtaaaacc | ctgtctctaa | aaaagaaggg | agttagggct | gggtgcgggtg | 200280 |
| gctcacggct | gtaatcccag | cacttttagga | ggccgaggct | gggtggatcac | ctgaggtcag | 200340 |
| gagttttaga | ccagcctgac | caatatgggtg | gaaccccgct | tctactaaaa | atacaaaaaa | 200400 |
| tagccgggtg | tggtggcggtg | cccgctgaat | ccagctact | tgggagggctg | agacagagga | 200460 |
| atacactgaa | ccagggaagt | ggagagtgca | gtgagccgag | atcataccac | tgcaactccag | 200520 |
| cctgggtgac | agagcaatac | tctgtctctaa | aacaaaaaaa | aagaaaaaaa | gaaaaaatttt | 200580 |
| ttaaaaatta | aattaaagta | aagggtgatta | taatcttatt | acaaaagtct | acaaaagtgtg | 200640 |
| acatccatc | acatttgctg | tattctactt | atcagaagca | agtcactaag | tcagccctac | 200700 |
| attcaaggaa | agggtagaca | tgggggtgac | gtccaggagg | cagggattac | tgggagccac | 200760 |
| gttaggagct | gtctactaca | agggaggaaaa | agcaatgaaa | ggatgaaatc | taggattgggtg | 200820 |
| gttatctctg | agtggaagca | agggggagag | atgcaagaag | cacatggccg | gtgaaagtta | 200880 |
| tttgtaataa | ttccattctt | ggggcaggtg | gtggatttat | aggtgttcat | aagtgagttg | 200940 |
| attatggaca | aatggggaga | gagtgctcatg | aaocaaaggt | tatatattagt | ccaattcagt | 201000 |
| ataatttaaa | atatccaata | ataaatgaaa | taaaaattgc | caactactcc | tgaccttggt | 201060 |
| ctgagcccca | tccatgcact | ccacctgttc | ttagtttaac | actggctccc | ctcgaccctc | 201120 |
| cacttgcgcc | aaactaaccg | ttagcccttc | acacacactg | acccttgcta | tagcaaccag | 201180 |
| atctctccat | agctcgcatt | agggcccccg | agttctattc | ctgaaactt | ttatccatgt | 201240 |
| aacctcactg | ctaggctgct | atagttagaa | ttgctgagaa | tggtctgatg | agtgggccga | 201300 |
| ggaccaggag | acaggggcagg | ctaacagttc | tgcccttagca | gtgacactgt | ggggccagaa | 201360 |
| gacaggtttg | atggctggta | tcacgacctc | acctattggc | ctcctacatg | gcaaaaggctg | 201420 |
| ggccagggtc | tgatgggtct | tcctctttga | ggctctctgt | cctggatatc | agatatagct | 201480 |
| aagtcgtgct | ctgagaagtg | cacctctgcc | acagcaccag | gcoctgtaact | cccttggttg | 201540 |
| taabtagact | tatcaactct | ctcctttcct | gtctttgatg | aagttaagcta | ccatactgga | 201600 |
| gatggctatg | tgctcaaggaa | atgagagtgg | ctattggcca | acaatgagta | aaacactgaa | 201660 |
| gctttcagtt | tgacagccca | ctaggaagta | aattcagtc | gcaaccaagt | agccctggaa | 201720 |
| gccttggatc | ttcctcagtt | gtgocctcag | atgagactgc | gttctctattg | gcctttgct | 201780 |
| tgcagccttg | taggaaacat | tgaagtggag | gacccagcca | agcctggcct | ggactcttga | 201840 |
| cccacagaaa | atgtgagata | taataataag | tatttgttgt | tctaagctgc | tacattgggtg | 201900 |
| ataatttgct | acacagtaag | agataattga | tacaggttct | tttcaattct | tttgttatta | 201960 |
| caaagagtgt | ttctctgggc | caggcaggtg | ggctcacgcc | tgtaatccca | gcaacttttg | 202020 |
| aggccaaggc | aggtggatca | ctttaggtca | ggagttcgag | gccacactag | ccaacatggt | 202080 |
| gaaacctgtc | tctacaaaaa | atataaaaaa | ttagccagat | gtgggtggcat | gcacctgtaa | 202140 |
| tctcacatac | ttgagaggct | caggcaggag | aatcgcttga | actcaggagg | tggaggttgc | 202200 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| agtgagccga | gatcgtgcaa | ttgcactcca | gacctggacga | cagagcaaga | ctccatttta | 202260 |
| aaaaaaaaaa | aaccaaacc | aaagagtgtt | gctatgaacg | tcgttgtaac | atgtctccca | 202320 |
| gaacacagaa | cacaagagcc | taggtttccc | ttgggtatgt | gctgggtata | tgccataggaa | 202380 |
| tggaattggt | agatctcagg | atgttttggtt | tgttttgaga | cagagtgtcc | ttctgttgcc | 202440 |
| taggctggag | tgcaagtgtg | cgatcttggt | tcactgcaac | ctctgcccc | cagcttcaag | 202500 |
| caattctcct | gcctcagctt | cccaagtagc | tgggactaca | ggtgcccggc | accacgcctg | 202560 |
| gctaatctgt | cgctattttt | atagagatgg | gttttcacca | ttggccaggc | tggtcttgaa | 202620 |
| ctcctggcct | caagtgacct | acctgccttg | gcctcccaaa | gtggtgggat | tacaggcctg | 202680 |
| agccaccatg | cccagctgga | tctcaggatt | tacaaatgtt | ccacattact | ggtagtgcga | 202740 |
| aaaccattaa | agttagggtc | atgccatggc | tgtagtgtgt | tgggcagggt | tgggagttaa | 202800 |
| gcctggcctc | aatccatgga | agttgggatt | gtctgtgggc | agggttgagg | tcgtacagag | 202860 |
| ggccagagta | gggtgcagtc | cttacgtagg | gttgagggtg | agtggtgggc | cacatttggt | 202920 |
| aacaaattga | agcttggttt | gggtgtgggt | ttggacaagg | ttggagaaca | ctgtgagggt | 202980 |
| gggtttgggg | gtcaatctaa | ggccagggtt | agaagttagg | ctgtcatctc | gtgcacaaac | 203040 |
| tgacccctcg | tgatatcatc | aacctaccac | gctgtggccg | caacaggacc | agccaccac | 203100 |
| aagatgagat | ccactctggg | tccagaaaag | tctttcactc | aaggcggggg | tggggagatg | 203160 |
| gaggacaagt | gaagaagaaa | cttggtcccc | cagatgcaa | aaagagaagg | caggcacagg | 203220 |
| aaatagaaca | caacactgac | tttaattggg | gagccctgag | ccgtaacctt | caaaccttcc | 203280 |
| gcctcagtac | ccctgtgcgc | aggagggagg | ggcgactgct | acgggacacat | cgtcagatgc | 203340 |
| ctccctactc | ctggcgatcc | cacgcctcac | cccttctctc | ccagctgctg | cgtgtctcgt | 203400 |
| cgccgaggtg | ggtccccagg | aagctccacg | gaggggcagg | ggagcccttc | agcgccccgg | 203460 |
| gacagatctc | cccatggctg | aggcgatggg | gaactagcct | ggatggagac | gcgcgcgtcc | 203520 |
| tcggagctgg | gcggggacct | gattctgggg | gtgtggacag | agccaaggga | ccgcccccca | 203580 |
| aggccagcgc | ccgggagatg | caaggccggg | cccaagggtg | tggaactctg | tttgaggggac | 203640 |
| aggctaggtc | tctgcacgtg | gcttctgggc | tctggaaaag | ggctccattt | ctgcacccgg | 203700 |
| atctccggag | tggtaggagg | cggtctcagt | ccgggcctgc | ctcttagagg | ttctgtctcc | 203760 |
| atctcctccc | acgtctaccc | atcccaagga | aggagggcac | tcgggcccc | gcaggctcgt | 203820 |
| gagagcagcg | ggctccgccc | tcccaatggt | ctatccatcg | gtgggtgggt | ccggccggct | 203880 |
| gtcggggctc | tgccgggtac | ccgcggctcc | ccgcggccgc | cggccgcgcg | ctcacagctg | 203940 |
| cttacgtccc | gcctgtctga | gcccgcggga | gaggtcttcg | atccgacatc | tcttgagctg | 204000 |
| cgacaggtgc | tccaggcgcc | taccttcat | ctgcaggacc | taaggccgac | cgccgaggcg | 204060 |
| tcaagcctgg | cggtctgtct | ctctcgtccc | ggcctctgct | ggcccgagcg | ccccaccgcg | 204120 |
| agccttcata | gaccctcac | ttgcacggtc | tcttgagagg | ccaacagctc | ctgctccttt | 204180 |
| tdagcgatct | ggaggacgaa | gctggggtcg | ccctgcacgc | cccggttata | cgtggaggcg | 204240 |
| cgaggcgccg | gcggccggct | ccagctgggg | tggggaagcag | cttagcggag | gcttggaact | 204300 |
| cgggccctac | actcacctct | ccagggcctc | agaccaggcc | ttcccttctc | ccctctcagc | 204360 |
| ccctttctcg | cctcccgctt | caggctcagc | gtaccgtgcc | ccccaccccc | agctgtgtgc | 204420 |
| agagatcacc | agcctgggtg | agtctggaac | tccagcttcc | tttacctgag | ctgacccctc | 204480 |
| cccggggggt | ccgggacacc | ttcacctcgg | gccttctggg | atacacctga | gacaaggcaa | 204540 |
| gtggagggaat | gactgtgctg | gagctcatt | ctcctgcttc | ccctccacta | cttccactct | 204600 |
| cacaagtgcac | aggaagaagg | acagaaggaa | agctccaacc | gggaggggag | aatggaaaaag | 204660 |
| gccaccttac | ccacatccat | gtagccactg | ccatcctggg | gagccagctc | ctgaaaaaac | 204720 |
| cccggggttc | cgctgtaagc | ccgggctcag | tcacccccaa | ggagccccat | gtcccttcca | 204780 |
| cgggccccca | ggccaggctc | tattagcaaa | gcacccctcc | ccccacacta | tcacccaggc | 204840 |
| tggcctttct | gtggtctggg | gagcgcccc | caccagcccc | ccagagcagt | gggagaagct | 204900 |
| ggagcggggg | ggcacctgta | aggagccggc | gccctgcttc | ctgcggcctc | gcctctctcc | 204960 |
| caggcgctgc | ctcagcggga | tgagcaccag | ctccaccacg | cctggggcgc | actgcgcgat | 205020 |
| cttgcgcata | acgtcatccg | gtactgcgaa | gttcagcctc | ttcagtaact | tctgaggggg | 205080 |
| tagacattgg | gatagcagat | tggttctctg | ggaaatgaag | cgaggggatg | ggggagaggg | 205140 |
| aaagggggcat | gggtggtgct | gttctgagtg | tacagccctg | tgggagttaa | agcgcctcgc | 205200 |
| atctgagctt | ctctgtgtgg | gttctctgag | gttctctgoc | gttctcttaa | tagaagagtc | 205260 |
| tcttgatctt | gtaatcggac | gaattctgtc | tgctgtgacc | ctccccaccc | tgccaccaaa | 205320 |
| acagctctcac | cctgggcaac | caagaccata | ggagatgatg | gggctgcaca | ggtattctgc | 205380 |
| tacctgttca | gatgaccoca | gttctgtagc | ttctgctgga | gagagtgtgc | ggggacataa | 205440 |
| tttgtcatct | ccaccatctt | gggggaagtaa | aacttgatga | cctctgcaac | aaggactcag | 205500 |
| ggagagggga | gcagacaagg | aaagtggggg | catgggtaag | gaaggaggag | gtgagagaag | 205560 |
| gtgggggtgg | gagaatcagg | catggcagag | tggtggaacc | caatggaggg | gtggagatgc | 205620 |
| ggaaagggga | aggataaggga | gggagacaaa | gacattaaga | ttagtgatag | gtagaggaga | 205680 |

66/154

| | | | | | | |
|-------------|-------------|-------------|-------------|------------|-------------|--------|
| ctaagtctcc | agcatacctc | agaggagctg | catcccttcc | ccactaattt | aggaatggga | 205740 |
| atactgagtc | tctggtcctc | cagtagagaa | tgagtcagcc | agtgcagggg | ctcacacctg | 205800 |
| taatcccaac | actttgggag | gctgagggcg | gtagatcacc | tgaggccagg | agttcagagac | 205860 |
| cagcctggcc | aacatggtg | aaccccatct | ctactaaaaa | tacaaaagtt | agccaggtgt | 205920 |
| ggtggtgcac | gtaaacccag | ctacttgggt | ggctgaggca | gaagaatcgc | ttgaactcag | 205980 |
| gaggtggagg | ttgcagtgag | ctgagatcat | gccactgcac | tccagcctgg | gcgacaagag | 206040 |
| tgagactctg | tctcaaaaaa | aaaaaaaaaa | aaaaaagaga | atgagtctcc | aagagaagct | 206100 |
| ttcaggagcc | cagggaatct | acatggatgc | atgatgcaca | cacacacaca | cacacacaca | 206160 |
| cacacacaca | ctccaacacc | aaagttccag | gagagtaaaa | gactcacgta | agcagagaca | 206220 |
| catgtgataa | gatcccccta | cacatagata | catagggcag | aatgtgtaga | cacatagaac | 206280 |
| aaacacatgc | tagacacaca | taggcacaca | tgtagaacac | acatagacag | aagcacaaaa | 206340 |
| agccacatat | gcacacagcc | tccccaaata | tacagatata | gagatgcata | gaagtacttg | 206400 |
| atcgaaaaag | atgacagata | agtattttctg | tcattgagat | ccaggagagc | attcctagaa | 206460 |
| acacaaaacac | cttcacaatc | acaaaagaaac | acatgcacac | aagaccccc | caaatacaag | 206520 |
| actaaacatg | tcaaacata | tgaacacaca | catgctcaga | ctcccagaag | acatagaaca | 206580 |
| catagcaatg | agtgcagaca | tacagaaatg | cataaatata | tatgcagaca | catatcttaa | 206640 |
| tcatttaattg | ccctggaggg | agtttagaggc | acccccccca | acatgcacac | aggatagaac | 206700 |
| atgtagacat | acatatatgg | acacacatcg | ttcacacaca | gatgcaata | gaaaacaagc | 206760 |
| acacacatct | aaaaatttgaa | gtccccaggag | gcagattcct | ccccactccc | acaagcacac | 206820 |
| aaacacaaaag | agtttgacac | gagacccgtt | acacacgaac | atacaggaca | gaacatgcag | 206880 |
| acacacaagg | gcacacacct | ccatcgctaa | agtcgccggga | gaggtttcgc | ttggggccggg | 206940 |
| acagagggat | gttgtctacc | cacagggtaca | gctggtgcag | cgctctctcg | tccacgctgc | 207000 |
| tcgcacctgg | cgctctcacg | gcttgccgcg | cccagcggtg | ccgggtcccc | cccagcctc | 207060 |
| acgacctttc | agggcgcttc | tctctcgcca | ctgcagaagc | ccataactgc | ctctgctctg | 207120 |
| ctaattccaga | gactcacgct | ccagctggga | gcggccatat | tgttttctga | aacctggaa | 207180 |
| accattcccc | actctcctgg | gtacaaaagac | ttctgggagt | tgtagtttta | tccccctccc | 207240 |
| gctacctcct | ctcaggcccg | agctttgaaa | aggcggttag | ctgcccttgt | cttcttccca | 207300 |
| agaaagcatc | acttctgtgc | ccacgcccac | tagtgaccag | ccaactcatc | attttgggac | 207360 |
| cagccaaaagc | caaccaggac | cccagaattc | cgcggatcct | tctatagtgt | cacctaaatg | 207420 |
| tcgacggcca | ggc | | | | | 207433 |

<210> 6
 <211> 23574
 <212> DNA
 <213> Homo sapiens

| | | | | | | |
|-------------|-------------|------------|------------|-------------|------------|------|
| <400> 6 | | | | | | |
| gctggacctt | caacagagaa | ggatggcttg | cagagtccta | gtctgatagg | cgcccccttc | 60 |
| ctagtgcctc | tggagcagga | gcacagtgaa | aacccagctg | tctgcgctga | tacccctggg | 120 |
| aggagccttt | gctgcagcaa | ctcccccgcc | tcatttaagt | tccaccccgag | gctgcaccgt | 180 |
| caggtaaaac | tgggagccac | gggatttcgg | gcctgaatgt | gcctagacgg | accccaacac | 240 |
| actagccctg | ccacgcacac | ccagcaacac | acagacaacc | acccagggag | acctggtctc | 300 |
| tcccatgcata | aactagggaac | gctatgaaac | agcatttttc | ttttcctttt | ttatataatt | 360 |
| ctatttttaaa | cttaaaaaaa | tttttttgaa | tagagttggg | gtctcgtctc | attgcgcagg | 420 |
| ctggtcttga | actcctgggc | tcaaaccatc | ctcccacctc | ggcctctcaa | agtgcctagg | 480 |
| tgacagggtg | gagccaccat | gcctggccta | tttttaacct | cttattgttg | actaatctta | 540 |
| gacttgcaga | aaagtctgaa | aatgtaagca | ttttccttac | acccttcacc | cagcttcccc | 600 |
| tgactggtcca | tcttacttaa | tcatagtcca | atcattacag | ctagggaaact | aaaccttggt | 660 |
| caatcctggt | aacgaaactg | tagactgttt | tgcatttctc | atttttccac | taatgtcctt | 720 |
| tttctgttcc | aagatccatt | ctaggtaccc | acatcacagt | tgtctcctta | ttctcctcaa | 780 |
| tctgggagtt | ccctcatctt | tccttatctt | tacccttgac | acttttgaag | aatcctgggc | 840 |
| agttattttg | cagaatgttt | ctcttgagtt | gtctgatggt | ttattctgag | cagaatgaga | 900 |
| cacagcatgt | ttttgactaa | ccaaaaagtt | attctataag | taaataattga | ggttaaaaaa | 960 |
| ctcatccaac | ctgggcaaca | gagtaaggcc | ctgtctcaaa | taaagtctca | acactaagat | 1020 |

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|------------|------|
| ttaaaaaagtg | accagaaaaag | ccccactat | gatttgctt | gacttttttt | taaaaaacaa | 1080 |
| acaacaacac | aaacaaacaa | acaaaaacga | tgtcttgctc | tctctctccc | tcaggctgga | 1140 |
| gtgcagtgcc | acgactcttg | ctcactgcac | cctccgcctc | ccagggttcaa | gcgattcttc | 1200 |
| tgcctcagtc | teccaggtag | ctgggattac | agacgcgccg | cccatgcgcc | ggctaatttt | 1260 |
| tgtattttta | gtagagacaa | ggtttcacca | tgttgccag | gctggctctg | aacttctaac | 1320 |
| ctcaagtgc | cgcgccactc | ctgcctccca | aagtgtgga | attacagacg | tgagcaactg | 1380 |
| cgcttggcct | cattagcabc | ttaaatctcc | acacagggtg | tggttccctta | ctgttataag | 1440 |
| gagcaaaagg | tcagtttgag | gacaggtaaa | ataaaaatgc | gcttgctgcc | tagagggaga | 1500 |
| agtcctctgt | gaagatagct | ttgcttgaat | gagctcaatt | gcaatgcag | tgctgaggct | 1560 |
| tgttgactgt | acggtcacca | cagttgctgc | tgcgcgccca | gaacatgggc | actttcttga | 1620 |
| ctacctaacc | tgtctcagta | catctgtctg | tggtttggtg | tggtccattt | cctaattttt | 1680 |
| ttaatgaatc | agaagactgt | gatgtgcttt | ccgctgtgct | aaccatggcc | gctgaagcaa | 1740 |
| aatgtaaaac | aagatgcgcc | tgaggtggtt | gtgcttcact | ctacgacatc | tgttaccgga | 1800 |
| aagggggtcca | gattcagacc | ccaggagagg | gttcttggtg | ctcgtgcaag | aaagaatttg | 1860 |
| agacggagtcc | ataaagttaa | agcacattta | ttaagaaagt | aaaagaataa | agctaatttg | 1920 |
| actccataga | gagcgcagcc | ctgagggtcg | ctgggtgccc | atttttatgg | ttgtttcttg | 1980 |
| atgatctgct | aaacagggtg | ggattgttca | tgtctccctt | ttttagacca | tatatggtaa | 2040 |
| cttctgata | ctgcatacgc | atctgttaac | tgtcatgggt | tgttggtggag | tgtagcagtg | 2100 |
| gggaccgacc | agaggtcact | ctcatcacca | tcttggtttg | ggtgggtttt | agccagcttc | 2160 |
| tatatgtcaa | gctgattttt | ttgggttggt | tgggttttga | gacggagctc | cgctccaggc | 2220 |
| tggagtgag | tgacatgac | tcagctcact | gcaacctcca | cctctgggtt | tcaggcgatt | 2280 |
| ctctctgctc | agcctcccaa | gtagctggga | ttacaggcac | acaccaccat | gcttggtcag | 2340 |
| tttttgtatt | tttagtagag | ataggttttc | accttggttg | ccaggctggg | ctgaacttc | 2400 |
| tgacctcagg | tgatccgccc | actcagcct | cccaaagtgc | tgggattaca | ggcgtagact | 2460 |
| accgcgcctg | gcctactgca | aactatttta | tcagcaaggt | ctttatgacc | tgtattccct | 2520 |
| atctcatcct | gtgacgcaga | atgctgttaac | tgtctggaaa | cgacgccag | taggtctcag | 2580 |
| ccttattttg | ctcagccctt | atccaggatg | gagttgctct | ggttcacagg | cctctgacac | 2640 |
| atctctctgt | gttttggcgt | gtgggaggaa | agaggggtga | gggaaggaaa | ctcaaaaaca | 2700 |
| agctctgaac | acacagctgc | ggtaacactc | cccactgtc | tgtgggtgcc | acaagtcaag | 2760 |
| ggagggggcag | agagagaaga | aggtgtgaca | gatggccgca | ggccacagaa | tgtcagagga | 2820 |
| agccaggttc | ctcccggggc | agcccaagta | gctggtagtt | gggtggcccaa | acagagggcg | 2880 |
| tcacagctga | gctggcgctg | ctcgttacc | gcagctcagc | gtccactctg | ccctcagta | 2940 |
| cctcctgctc | agcctcaggg | tccactgccta | cctcctgct | tccagtcac | ttctgctgta | 3000 |
| ctcctgcttt | tctgctgttg | gccccactgc | agctcctttc | tgtgagctt | ctcttctcca | 3060 |
| gttctgcagc | acagccaggt | gatcctgggc | tcacagacag | ctctccccc | agtcctgggc | 3120 |
| ctccctctct | agagccctct | ttccttccca | cgtggcctcc | ccagggttgc | ccactgaatt | 3180 |
| gagaaggggt | ggagggggtg | ctgggcagtc | ttggggatta | gccaaggagg | cagagttggc | 3240 |
| ctccccaggg | tcccttgtag | ctggagtcct | gcggggtcta | gacacccctt | cctgaagggt | 3300 |
| aagagcgggg | gaggtatatt | aacgtgtatt | tttagagctt | ctcttttttt | tttttttttt | 3360 |
| tttgagacgg | agtcttgctc | ttgttgccca | ggctggagtg | cagtggtcag | atctcagctc | 3420 |
| actgcaacct | ccacctccca | tgttcaaaaa | attctcctgc | cgcagcctcc | cgagtagctg | 3480 |
| gtattatag | cagctgccc | cacacccggc | taatttttgt | aattttagta | gagatggggt | 3540 |
| tgcataccat | tggccaggct | ttctgcgaac | ttctgacctc | aaatgatcct | ccgctctgg | 3600 |
| cctcccaaa | tgtctgggatt | acaggcgtga | gccaccatgc | ctggccttag | attctctatt | 3660 |
| ttatgatggt | ttaacatctc | gggttggggg | ctgttggtgc | ggagagaaac | tgttgatttc | 3720 |
| ctggagatca | gaaacaactc | atgcctttca | tatgcaaac | gaccagctct | gagtcctaac | 3780 |
| accaaccacc | cccttcaagg | aactctcaca | tacgaaacca | gtatttcccc | tgccttaaac | 3840 |
| cagctcaggg | ccaggcaccct | acccagacaa | ttagagccca | acccacggcc | ccaaacccca | 3900 |
| ccagaatgat | tcaaaatgcc | aatcctaccc | tcttccctcg | gcctgccttg | cctcccagct | 3960 |
| ggaaaacggca | ctgtgggctg | tggcctgtgc | cttccactcg | ctcctgattc | tgtccctgga | 4020 |
| ccaaacctag | tgcctcccca | ctgtggcctc | gcattggtgg | taataaacta | taataaacta | 4080 |
| tttcaacagc | attggcctct | tgtgcatcag | tcaccttcat | aaattaaaat | tttgcaacta | 4140 |
| caactgaggg | aggagggggc | tttagagagc | actctcacct | ggccctcttg | gcttgtggag | 4200 |
| tggagccgca | ggtagaagttg | cctcctcgca | ctcagttttg | ggatgggttt | gttatggcg | 4260 |
| aacagctggc | tgatataagg | cactcagcag | ctcttgcatg | aagcaatggg | agaattgtga | 4320 |
| gcgcccaagg | agcaggagt | gacagagcaa | agaggggtgt | caaaactagg | aacacgctcc | 4380 |
| tgtgcccaga | cacatgcctg | gggctctggc | tgcaccttag | gctcggcctg | ccagccacca | 4440 |
| cccgctcagc | caccacccca | ccaaccacca | ccccgccaac | caccaccccg | tgtggctctc | 4500 |

| | | | | | | |
|-------------|------------|------------|-------------|-------------|-------------|------|
| agggcacc | atgggttgaa | tttacatata | gaaaagacta | accaaggtcc | aagttaatat | 4560 |
| gtctttttaa | aatttatatt | agagacaggg | tcttgctctg | tcaccagggc | tggagtgcag | 4620 |
| tggtgccacc | atagctcact | gcagctgttg | actcagggtc | aagtgaatcc | ctctgctcag | 4680 |
| cctctaggag | acagtagcct | acaggtaac | accaccacgc | ccagctaatt | ttaaaatttt | 4740 |
| tctgtagatg | cggtgtcttg | ctgtgttgcc | cagggtcgtg | tggacctgct | ggcctctggt | 4800 |
| gatoctccta | ccccggcttc | ccaaagtgtc | gggattatag | gcattagacca | ccacgcctat | 4860 |
| agccaacatg | tctttctttt | gactctact | ttggtatctt | ttctttaaag | gttccctctg | 4920 |
| tcccccgac | acacacagaa | tgggggagag | gctgtcagat | tctgagctcc | agaacctcag | 4980 |
| gtgtgcact | gggatgtggg | gtgggggctc | aggaaccacc | taggggagaa | gacagggttg | 5040 |
| gaagaaacag | gaaggaaggt | ccccaaaatt | atgtttgttt | gcagaggcca | gccaggctgc | 5100 |
| aggggagtg | ggactcagtc | gaacataggg | gccccaggac | cactagcttc | tggccagcag | 5160 |
| tcactgccctc | cacagagctg | ggtcgttgga | aattgcatgt | aggagacaca | ccagactccc | 5220 |
| aggacagagc | ccttttgsga | tggccagcac | taccagcct | ccactggtga | gggaggtcag | 5280 |
| gggtctgtgt | acccttgctt | ctgggactga | tggtttattg | agctggagag | tgtgccaccg | 5340 |
| agtgttctcc | agccctcagg | aactctctaa | gtggctctgg | gttccctggag | tgggtgggtc | 5400 |
| gaagctccac | tccgggaaga | aacttccaag | ctgcctgcag | gtgctggagg | tccggtgatt | 5460 |
| cactggctct | gcccctgcag | ttcaagttcc | tggagtggct | gtcagtgggc | acctgtcttt | 5520 |
| aaatgtgttc | atcttaggag | ctactctca | ccagaggcag | gactctggca | gtctggacttg | 5580 |
| atctgctgag | aatgaggagg | atatgttgtc | ccctaaggac | tggggcccca | ggctgcaagc | 5640 |
| tgtgtggcag | agagcccatc | ctcactcagt | gaggaccagt | gatccaggaa | aagccacagc | 5700 |
| ttctccctcc | ccagcccagg | ggcttccagc | atcctgggtc | ccatgataac | caagaggtca | 5760 |
| taaaactcatt | tccataataa | ctctagccca | gaaacctgat | tagggggcag | caaaactagg | 5820 |
| gggtgggagag | gtgggagggt | gggcgatgag | agggggaggc | tttgaatcca | ggtccctggc | 5880 |
| taccttgggg | gtcaggcgag | actgctggca | gaggcttctc | agggtggctg | ctgggctcat | 5940 |
| gagagttctc | aggggtctgg | agaaatgggt | gagggtaaat | gttgtgaata | tgtgcagcag | 6000 |
| gagacctggg | ggctggggag | gggcataagg | gactcaaggt | gactgggtgc | tgcccatctg | 6060 |
| gaaggaggcca | ggaggcatga | gcccttccct | tctcccttcc | ctctccacct | cccctgggtg | 6120 |
| cctcactcac | ccaggggcca | gggctgtcca | gtggctgttg | ggcccaactc | catgggggtga | 6180 |
| acgcgcgcca | gggggtgggt | cctgtgttg | gggtgggtgc | ccatctttgg | ctgtgataaga | 6240 |
| gtccaggagg | ttggcacagt | gatcctgagt | gggttatgtc | ctcccgcagc | catgggtgtcc | 6300 |
| agcccaggga | gttctcgctt | tactcagttt | cttggggcac | ccatctgtct | caagtcaacc | 6360 |
| tctcagctcc | cttctgtctc | cctcttcagg | ggagccttgg | gatccaggct | ccaagttagc | 6420 |
| ctcactgcct | cggctggcac | ctcctctctc | tagtccctaac | atcttctcca | ggctctgaca | 6480 |
| ccaccacgca | gcctggcact | ctccagatgc | tggcatcgct | cagctctcca | agaaaccttg | 6540 |
| atgtccgccc | cttcggcagc | tatgtctgct | ctccttgccc | ctgggtgccc | tgtgccectt | 6600 |
| gatgattcca | agccatcttt | gactgtcccc | atcccattccc | ccaaggcctt | gtcatttccc | 6660 |
| gtgatgttcc | ttcaaaacat | tctccctgc | cctgagactc | ccgcttgggg | atgagaagca | 6720 |
| gccgccaccc | tctgcagcgc | cccctccctg | ctgacacgcc | aggctctggc | caccttgctc | 6780 |
| ctctgccacc | agacccctca | tcacaactcg | ctttgtcagg | gtcctcttag | ctgccaccgc | 6840 |
| gggcccagggt | gggtcccttg | cctctctgt | tatgcccctc | gcccctatct | ctggcccaaa | 6900 |
| tcactgccatc | tcccttggtc | tggccggagc | actcccaaga | ccaggctatg | tcagacatgc | 6960 |
| ccacagagtg | ctgcctctgc | ctagggccct | ggtgcagggt | gagtcctagg | acagccatgc | 7020 |
| ttagtattat | gtgactcccc | actcgcgcac | cacccaggtc | acagagaact | gggttaaggc | 7080 |
| agggccctgg | cacaggggca | gccagcaccc | cagctgacca | gtggatgga | gtgaaaagat | 7140 |
| gtgctgggccc | cagcatttgg | gaacttcaag | ggggtgacag | aggtgatatt | tcagagggaa | 7200 |
| gtggcaaaag | gccgaaaaat | ggtgagacag | aggctggaca | ggcctccggg | ggcagcatgg | 7260 |
| tacagggaact | gcaatctgag | ccaggggaaa | acaggggcga | gtcaagggtg | aggcagccag | 7320 |
| ctggtgggag | aagcaggaga | gtggacaaga | ggagctgtac | tgggaggtag | agggccatgc | 7380 |
| cttgccgtgic | tgggtgggtg | gcagggatcc | acccccctgc | ttgactggga | ggccacttga | 7440 |
| acctcctgtt | caaagctact | tctttccatg | gcctctgggg | ctgctctctg | cacctggggg | 7500 |
| caaggctgag | ggcctgcctc | agctcccaca | gccccagcag | agctctgagg | aggggaaccg | 7560 |
| caggagttag | ctcaggaaag | aggcgtctcg | agcctaccca | ctgcacgcag | ggtccctctc | 7620 |
| gcagcccccag | ctgactcgct | gcagatgggc | tcttgggagt | cggtagcaca | accagccatg | 7680 |
| gccggccccc | gggagcagag | gcagcaggac | gctgaggagc | atggccagca | ggaaggtgtc | 7740 |
| atggtctcgg | gggatttggg | gaaggggcgc | tgagtcctga | gcaggtgcac | caccccagct | 7800 |
| cctgccccaca | tgcccccac | tgcatactt | tgcagctgca | caggggccat | gtccatgctg | 7860 |
| ccaccaaagc | ctggcttgtc | acagaaggg | ggagcccagc | ctggagcaca | gtggcagtta | 7920 |
| tggttgctat | tgcaaaacct | cagagaagag | aagaggaggg | tcacgtaggaa | ttaggaaccc | 7980 |

| | | | | | | |
|-------------|-------------|--------------|-------------|-------------|-------------|-------|
| caaggtcacc | cccactcctc | gggctctcac | cccgtaggctg | tggcaggcag | tcaggcagcg | 8040 |
| ctgaagctcc | tgggaaggcat | tcttctctgca | gcgcctgctc | tggcacacct | acggggcagtg | 8100 |
| caccaggcag | tgaggggggac | actggcctgctc | gggattcaaaa | cggcaaggag | gggtcgggtg | 8160 |
| ggcagagctc | accattctag | gtcccacactg | ggtgcctggc | tctaccaggc | ccaggccaag | 8220 |
| caggctccagc | tgggcactgg | ggagtgccaa | ggctccccga | caagtcactt | cctggccatc | 8280 |
| taggtgaagc | gtagagtcca | ctgggaccat | gtgcggtgctg | agcaggctgg | gctttccacc | 8340 |
| ctggcactgc | agcttccac | acaggggcatc | cctggggagg | aagtagagg | gggtcaacag | 8400 |
| ctgcagtacc | ccctctcccc | aaacccactc | catagcttct | gctccctcca | ctcagctcca | 8460 |
| ctccctacct | ccctgcacag | ggcagggaagt | ggccctcgct | gtcctggccg | cagtttccat | 8520 |
| gagcatctcc | cgcagagttc | accacotgga | aacaggcctc | gggagctggg | tgggagcctg | 8580 |
| aggaaagcatg | ggccagggtg | ggggcagctc | ggagaggggc | tgcgctcagc | gggcctcagt | 8640 |
| ttccctggcc | catcttcccc | acagtagaaa | actggcccca | ccagaggatg | gggggcaggg | 8700 |
| tgaagggtg | ctcgtgtcct | ctcaccaggc | ccccagagct | gctggcactg | ctgctccagc | 8760 |
| gtggggacatg | cgccatccca | gcagtagcca | ctgcccctgg | cacagggtga | gccgtccagt | 8820 |
| aggtaaacgt | ctgggggaca | gtgcgcgtgc | gtgcccgtgc | aaaaactcag | gaggtcacag | 8880 |
| tcacccatgg | cctggcggca | cagcgctcca | gccggcttca | gctgcgcagg | tgacgggtgg | 8940 |
| tgggggaaggc | agagagaggc | cacgtgcagt | gagaggtcca | tgcagagagc | gcggctcgga | 9000 |
| gctgggggag | ccaggcctac | ccaagcccag | cacccaacgg | gggaaactga | gggtaccaat | 9060 |
| taactaaggc | caacaggccg | gctcccaagc | tccccgaaac | cctcacctcg | aaccttccat | 9120 |
| ggcctcacca | ggcagcgcac | gcagcagttc | ccgtggggcg | actggggccc | cggggcgagc | 9180 |
| gagcagttgt | gagcaaaagca | gcagaggtcg | cggcactcct | gggaccagaa | aggcaagaag | 9240 |
| ggcccagggtg | agggcgagc | gccccagacc | tgagcggaga | gggcaagtga | gggcccgggt | 9300 |
| agccagctta | acctggccag | ggcgcgagtc | acactcctcg | cccgcttcca | cgaagccggt | 9360 |
| cccgcagagc | gcccggggca | gggggagttc | ggggctccgg | gcattggaga | ggcaagcgcc | 9420 |
| gccccctctg | cgggaagaagg | ggcgagctg | gcccgggtcg | caggcgctga | acacgcgcgg | 9480 |
| aaacgggtgc | ctaccggcac | ggggaggcca | ttgggcatgg | agggacagtc | ccccaccccc | 9540 |
| cgcgcttctc | tgatccccac | ccctgggctt | ggctacagcc | gccagacgcg | cagagcccag | 9600 |
| agaggggaag | taacccgcgc | aaagtcacac | aacaagcggg | acaggggagc | atcgcgctcc | 9660 |
| aatagtgaag | agcccgggac | ccaaggtgga | atcgcgaccc | gacggtgctc | ctccgggtcg | 9720 |
| aggagtaaac | tcgcccaggtt | actcggaaaa | taatcttcat | accgttgaga | atccactttg | 9780 |
| cctgagcttc | ttccctttaa | gcctcataaa | ccaccctgaa | gcggacacta | tgatcattat | 9840 |
| ccccatttta | cagaagaggga | aactcaggga | cgaccaaaag | aacgcagcgg | aggaaagtcct | 9900 |
| caggactagc | cgcccccgcg | cagccccgac | ccccaccocg | cgtaccgggt | ggccgcagcc | 9960 |
| atgacgcagc | ctccgcagtc | ggccgcagcc | tccacgcagc | agccgtcggt | gtcgtggctg | 10020 |
| aggccgaggc | tgtggccgat | ctcatggggc | atggtggctg | cggcgccgat | gggggagctcc | 10080 |
| gagtggtcct | ggggggccgt | gggaggggcg | tcactgcggc | cgtagagcct | cctgtctctc | 10140 |
| cctcgccccc | gcccgcgggg | ctcacccgtc | tcaecgctcc | cgagctctcg | gcgcggcaca | 10200 |
| tgccctcgac | gggcgcagg | cccactgtgg | cgccctggaa | ggcgcggccc | ctggggggcg | 10260 |
| agcgcggcgt | gaccaggcgg | ggccggggagg | tgaggccgcc | ccaccgcggga | ccgcgctccg | 10320 |
| ggtcacaggg | accacagtgga | gcagctgccc | ggagtctgtg | ggccgctgcg | ccacacagccc | 10380 |
| ccggcgccac | tgcaggaaagg | cccagagcgt | ggcgttgccg | tctctcgctga | cgcggctcgag | 10440 |
| gtcccgctgc | gtccacacct | ctgaagccgt | cagcgccacc | tgaatgtcca | gagtcctgag | 10500 |
| aagctgaggg | cgagggcggg | gcaggccggg | gacagggcgc | cgactgcgcg | cgctggctct | 10560 |
| tcgttggggg | cccttctctc | tccccaaccc | ccaccagcac | ctgccttgctc | tgcgcgcgcg | 10620 |
| acccccatca | cgcctctctc | ccgcgcgcc | ccaacctggt | ccacgtagtt | ggcgacttcc | 10680 |
| aggagacgct | gtttgggtgtg | gttcaagttt | cggtgcccag | tcaagaactg | ggaaggcaga | 10740 |
| aatcccgggtg | gcttgagggg | ctgagctggc | cccatccctg | accccgccaa | ccctgggggt | 10800 |
| ctctcctcac | cagggtgtgg | tctgcacaa | tgtacagttc | caggctcttc | cggtcctctg | 10860 |
| gcgcttctcg | cctgocctgc | ggaggtgcaa | atggggaccc | tgagtggaa | ctgctgggct | 10920 |
| tgagccctga | cccccaaccc | cagctccccg | aagggaagttt | aacatgtttt | ctggaaactg | 10980 |
| ttctctcaga | cttcaataaa | atacttggga | ctcgaggcct | gtgaattcct | gtctctctgt | 11040 |
| atttgagggg | ctatagatac | agcattccca | ctcccatccg | atcgatgccc | ctgacctctg | 11100 |
| tctggggacc | accagggaagg | ctggtcatgc | ccgctttgtt | cccaggatcc | ctgtggccac | 11160 |
| aggttctctt | ccaggtgagc | agctgtctcca | tccgaaagat | ctcgtggggt | gagaagtcct | 11220 |
| tggagccccg | gggtggccag | ggacgcagat | aatagctggc | atctcctgct | agggtgatca | 11280 |
| ggccactagg | gtgcagagg | gtgtagcgg | gtgtgaggga | gctcttctcc | catcccaggc | 11340 |
| ccagcctcct | ctcccagagc | tcacctcatc | ccagagcagg | tgcagaggac | tacccaggag | 11400 |
| tcggggaagc | cccttactcg | cccttggttg | tggcaatgat | cctaggggag | aaggggcccag | 11460 |

| | | | | | | |
|-------------|-------------|--------------|-------------|--------------|-------------|-------|
| ccccaaatct | cagccaggggc | tggagcaaga | ggggcaagag | ggaggggtgtg | gtagggggtg | 11520 |
| gctccaaaccg | ccccttagga | atgcaaggag | gagtaggggt | aggaaatgggtg | gggggggtacc | 11580 |
| tctggcggtg | catcccagag | cccatgggaag | catctcaccg | tgtgggttggg | ggccagcacc | 11640 |
| actggctggcc | catctggggc | gtagtggggtt | tctatgtatc | ctggggccag | cagcctgctg | 11700 |
| agaggggggtg | ttacagggaa | caactgaattc | agcttctctcc | tgctctctcc | aggatgtctc | 11760 |
| ccagccttcc | tccctaaatg | ctaattggagc | agctttatga | gtgagacact | cacagtgtgt | 11820 |
| cttaggggaag | ggacaggagc | aatggtgact | tgctcagatc | agaaactctt | ggggctagag | 11880 |
| gaaggagcct | tggtgatggc | ttagtltgtgg | gagatgtgaa | batgggaagc | accagggagg | 11940 |
| acgcggggga | ggagtgggaa | taggggaaga | gtttgtgggt | cccaggggac | ctgcagcagg | 12000 |
| cagcaggatc | cacaggatcg | ggaggggagg | agtcaaggaga | cactgccgaa | gaatgggact | 12060 |
| tggagtltggg | gaaatgcggt | gacctcccc | agttccccctg | cctgctgccc | tcctttgtttg | 12120 |
| ggcactctggt | cgacctctct | gccccacct | gccttagatc | cttgaaatat | tttctcaga | 12180 |
| cttctagacc | ccacatacct | cccacctgtc | cttcagtgtat | tgaatgtcac | cccctgcctc | 12240 |
| cagagaaaac | agaatcgcca | cctgccccacg | ctgcttccac | cctccccctc | ttctccacac | 12300 |
| ccactccggt | aatgattcca | tcttcaggct | ccatctcaac | aggatcttct | ccacacgact | 12360 |
| ggatcaatca | taagtcaatc | tgtctctctt | aaagaaaatc | cttaaccctc | cctcaccctg | 12420 |
| ggctcattac | ctccagacca | cccgtctaattg | atgggtctgct | ccccctctcc | aggcattcca | 12480 |
| ccactctgcc | cagctctgcc | ccctaccct | gccccacaca | cacccgcac | ctcaggaggt | 12540 |
| aggtgatgtg | accaccccat | tgaaagggtta | gggacgtcgg | gaaatatatg | ttgggacacg | 12600 |
| tggaaactaga | gtttgttccc | tgtccatccg | actccacgag | ggagaataaa | atacgtgtca | 12660 |
| agtgctcaga | acagcgctcg | cggtcaagca | ctcagtaggt | gatataact | gataaacataa | 12720 |
| tctgggtggt | tttaagagcc | tgcgctccag | ccgggacacc | cacccccctc | agccccaggc | 12780 |
| accttctcga | gcacaggtct | gcctgtccct | tccccagctc | aaaaatcttg | gctctggagt | 12840 |
| gtccaggaca | ctgagcacca | aaatggcctt | ctgtgatctg | gccccctctg | gactctccaa | 12900 |
| attcattccc | ccccgtctcc | ctccctgttag | atggagacct | tcaggcgagg | tcagacaaac | 12960 |
| tgtctgcccc | aaagtgtagc | tttctttttt | tttctttttt | cttttcttct | tttttttttt | 13020 |
| tctttttctt | ttttttctct | ttttttgaga | cggagtctca | ctctttgccc | aggccggagt | 13080 |
| gggtcagtg | tgtgatcttg | gctcacaca | acctctacct | ccgggggttca | aggcattctc | 13140 |
| ccgtctcagc | ccctcaagta | caggcgccgc | cacccagcct | gagcagcctg | ggctaatttt | 13200 |
| ttgtattttt | agttagagacg | gggtttccacc | atgttggcaa | ggctggtctc | aaactcctga | 13260 |
| cctcaggtga | gcgcccgcgc | tcgggcctccc | aaagccaccg | catctttggtc | cctgcattc | 13320 |
| ccctagcctg | gggtgcccgt | tcactcttttc | ctcttaggat | ctcttttagac | tcagcatatc | 13380 |
| ttgcaaatgt | ccactaggtg | gtgctcactc | atcgccagca | gggagctaac | aagccgctcc | 13440 |
| tggggcttggg | aggcggaagg | tgccccacag | cggggctgac | agcctcagcg | gtcctctcca | 13500 |
| gcctccagg | agccaaccac | aggcctgcgt | gactctccct | gtcatctgca | ccctctctgg | 13560 |
| ggctcctcgc | ccatccagcc | accccacacg | atctgtgtca | gtccccctcc | cccaacactg | 13620 |
| atccccctct | cccagcccta | cccagcctg | gcactcactg | gttcttctcc | agctcaagca | 13680 |
| ggagctcctg | gccttcagcc | tcacgggcca | ccagcccat | gtctggcttc | gagacctggg | 13740 |
| caagaaaag | tgtggagctg | agatgggtgg | ctccaggcct | cctgcccctg | agggagtagg | 13800 |
| tggcctgtgg | agccggctgg | ggaggaaagt | cttggggaga | acgtgggctg | gggagtcagc | 13860 |
| aggaccccc | acatactatg | gagggcgtgg | aggaggtgag | aaacatacaa | gatgttccca | 13920 |
| aaactcaggt | gattgcagtc | ctgacaaacag | ccacttggaa | gggcgttggc | acagcctggc | 13980 |
| aggcacacca | gcattcctccc | tagagacag | aggctccaga | aaaggtgcc | tcctctggcc | 14040 |
| ccgcctctct | ctttcatgcc | cagaaggggc | atcaaaagca | ggggaagaca | gaggggtgct | 14100 |
| gaggacatta | tgggggcctc | gggtagccat | ggtcaggggc | tcctcagagc | ttctgtctacc | 14160 |
| tgaggcttgg | ttccaaatga | gtgctgtctc | atttctctata | gaattcaaat | ttgactcctc | 14220 |
| cacttccaat | tttggcaaac | tgctcctctc | tcctaaagttt | tcctgggctc | ccagcagccc | 14280 |
| ccgttccctcc | ggctccgaca | cctgtctcac | tggacccacg | aagtgaacat | ggagccattt | 14340 |
| ccagccaaga | gagcacactg | gctctcagct | agggtgcagg | aggctggctt | ggaaggccag | 14400 |
| ccctctctcc | ttcccccc | ctcttggcgt | ctccccacct | gtgggaacac | cccacttccc | 14460 |
| cctgtccac | tcagcctggc | tggggggcca | gagttggagc | cggcccgagg | gcttctggg | 14520 |
| aggctgtcgc | gccttcggaa | tgtttaaccc | ccgactcctt | ttctccaaaa | atgcactggc | 14580 |
| ctggggccct | gtccaaggct | ctcagagctc | ttggagggag | ttcttctctc | gcaagtggg | 14640 |
| agcagtggt | ccttgctctc | ctggccacag | gccccacaag | gctctcagca | tggactcatg | 14700 |
| aggctggaat | gccacttgct | ttattgggga | aaggctctga | ccgggaaaaa | ggccatactc | 14760 |
| gagggtccctg | ttctctccctg | gcctctgcta | tctttactct | gtccccctctg | gtaccctggc | 14820 |
| ccttgatata | taccctctcat | cttgaaatgt | gagtggttcc | tgcttcttgg | aggggatacc | 14880 |
| tagcctctac | tcttctctct | gtaccattct | ggcaggcttc | ctgggggacg | gggcccaccg | 14940 |

| | | | | | | |
|-------------|------------|-------------|-------------|-------------|-------------|-------|
| gtggggggaag | cagagcccct | ttggggctct | cctcttggtc | acagcccagg | ccagacagac | 15000 |
| agggaggccc | agaggcagag | tgaccccagt | gtgtgtccag | ccttccccctc | ctggggatgg | 15060 |
| ggaggggcaat | ctcaaaagct | agggcagtg | cgtgcttgac | cagtggaatc | ggggcccttat | 15120 |
| gggcctaggg | gatcccagtg | agggccctgg | gttgggagct | gttgggtctc | tggggggcctc | 15180 |
| tcagccttca | tggcaatgct | ccccctgctt | ccctcttgct | ggatttggac | agtagggctg | 15240 |
| aaaattccaa | acaaagagg | ctctctagga | ggggcagggg | tgtagccaat | ggtttaaaat | 15300 |
| cgttcagacc | ttagtgggtc | tcaggctccc | agcctaaga | gctgtgtgac | catggacaat | 15360 |
| ttccccaaag | tctctgggct | tcctgttgcc | cctctgtaaa | atgagcatat | caaggctact | 15420 |
| gcccctcttag | tttgcgacac | agatattatg | gcacaaacag | atggggcatg | gttattcttg | 15480 |
| aagcgtgtga | agagcgggat | tgggaagagg | ctggggcaga | gcgtcctgca | gaagaagcac | 15540 |
| atgggggtggt | cttacatctg | ggggacatca | ggagagtgac | cactgcctcc | cccatccag | 15600 |
| aagtggattc | cacaggagcc | agtgggctg | aaggttcagg | ccttcgtggc | agggccctga | 15660 |
| gagggcagac | agtgtgtcca | cagggtccaca | tgttctggtc | aactttgcaa | aagggtttct | 15720 |
| tttttggtct | tttttttttt | tttttttttt | tagaggctcc | tgaaaagctt | caggaccctac | 15780 |
| aaactctgga | cccatttctg | cctggctggg | gtgggggtgg | cccagatcat | ccagggaagg | 15840 |
| agggaaagag | ggaggtgggg | tggagaaagc | tgaatgact | tccatgtgtg | cgggtcacg | 15900 |
| agatccagat | gtccaaaccc | cagtgctctc | ttctgcccac | tggaggggcg | ggggaggcag | 15960 |
| gggcctatag | gagtagtgac | tgggtgggtc | tggggacccc | agcaaaacta | gaagctgtaa | 16020 |
| tgtagggaga | gacaaaagg | ctgggagggt | caggggccct | gtggaggggc | gggagacatg | 16080 |
| gcactgacgg | gctcctccag | gctgacggtg | cgccagggtt | gtccatccag | gacccagtgc | 16140 |
| ggggtgactg | gctgcctcag | gagtagtctc | ggagtaaaag | cagagcacag | ggttgagggg | 16200 |
| acctgaggaa | cacaggggca | tgggacaaag | cagaggggag | gggttagagg | acatccccag | 16260 |
| ggaggcactg | gagggtcttt | ggggcgagct | tcaccttcaa | cacgcgtggg | ctcagcctbg | 16320 |
| agaaggagg | acgcccgtgg | gcaccccttg | atctgaggag | ctatcaagg | ggaggaaaag | 16380 |
| agaagcttgg | aaagggacag | ctcagctggg | gacacgggag | tccctcgacc | tttgtcgggg | 16440 |
| ggcaggcttg | ggctcggcga | tcacaaggaa | gaggccaaag | ccgcagctgc | agagggaag | 16500 |
| gaaagagcgc | ggcagcctta | gggattttta | gatgggcagc | agatgccttt | agggtagag | 16560 |
| atgtacgaag | aggagacact | tgtgcctccc | ccatcatctg | agaaaacaaa | cagccagatg | 16620 |
| ttgccttgcc | aggtccacct | tgccagagc | tcctcgggg | actctgtcct | ggtggcagg | 16680 |
| ttttggtacc | ctggcccaga | agggccctcc | tcactctctc | aaagggaagg | gacgtctccg | 16740 |
| gacgggacct | tgggtctccc | tggccgggtg | tgcctaagg | ggctctagga | ggaatccag | 16800 |
| agccaagcat | tactcagctg | ggccttgtaa | tgttccccctg | gaatgtctcc | agccctcca | 16860 |
| ctggcccocaa | cctcctcacc | agggccgccc | tgcaggagcc | agggcccagg | cacccagagc | 16920 |
| ctgcagcagc | ccctctctcc | ccgggtacc | gtccagctc | ccagaaacaga | gacgtccccc | 16980 |
| cctccacgca | gcccctggct | cagtcctgct | gggctgatgg | ctgcctgtgg | aagtgactca | 17040 |
| gctcctgcta | ggccaaccca | actccttttt | tctctccac | cttctctccc | agactacaaa | 17100 |
| catcaaaagac | ccttccctca | agaagccctc | cttgattgga | tgagtgaatt | gcatcaggc | 17160 |
| agatgagggc | cgagaggagt | ctgccacctt | ggaaaggagg | ctagaggggc | cagtgacagg | 17220 |
| agggctctga | gtggatgtgg | gggaggggaa | ggaggggagg | tctctcagcc | cagagagcac | 17280 |
| ttaactaga | gtagagaacc | aagctttgct | gctcctaggg | ctctaagggt | ttgggggaag | 17340 |
| ggtagggtgg | gcccgggcac | aggtgtggtg | tgggtgcagt | gtggtgtgtg | ggtgctgtcc | 17400 |
| acatggcctt | gcgcgcagct | ggtggccacg | ggcacccctga | ccccaatgag | ggagagagg | 17460 |
| cgagagctgg | agctggagct | ggagctccgg | tgaccgggtg | aattgggggtg | gaaccgagg | 17520 |
| gagccaggct | ggatltgggg | acatagagcg | ccctctccca | gggggtcccat | caactccctc | 17580 |
| gacccagga | tagggctcag | tggggaggga | cgagtggaac | gctcgtgggc | ctccccctgg | 17640 |
| gccagaacag | accaggcccc | tgtacctgtt | tgggtccccc | acagtgctgt | ggaagccacc | 17700 |
| gccagctctg | catagcacag | ccagcccccg | catgccccct | ccctgtgttg | cttccccctg | 17760 |
| ccgggccaag | cacttgctgt | cgaggactgg | ctaatctctc | cacccgcttg | cagaggttgt | 17820 |
| tccagcccca | tcttaacatc | tttgtttgga | ggggttaccc | cgaggagaca | gctgcagttc | 17880 |
| ttccagagca | ctgctaaaca | gacaccttct | atctggagag | gcccctctct | atctcaccac | 17940 |
| acaaggcaac | aatataaaca | acataacac | tgcccctgctg | ccctgggagg | atgggacgaat | 18000 |
| ggtgagcagg | gtggaggcca | cagctagttc | tgcagcctga | gagcacaagca | gggactctgg | 18060 |
| gggactcttg | ggcatggggg | cttctatag | gatggagccc | cgctgagtc | taaggggtgg | 18120 |
| aggagcagga | gcgggtcaca | cggtggcctg | cggatggaa | ctgggtgtga | gagcgagaat | 18180 |
| ccaggcagag | ggggctacgg | ctatgggctg | ggggctgggg | gctgggctgt | ccccgaggg | 18240 |
| ggggagccta | tgcctctctg | tttgccagc | gagtggcagc | cgggcaagtcc | ggcgaagtcc | 18300 |
| ggggccgggg | ccagcccaag | cacacttgag | cgtccctggg | caggtccccc | ggagaccccc | 18360 |
| ccaaagagtc | cccacgcctt | gaactactgg | cagtatgggtg | ccggggccggt | gagacccctc | 18420 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gcgcgctgac | ccgagctctg | agcagaaccc | atccccgcca | ccaccaccgc | gcctagcctg | 18480 |
| ccctcaggg | cgcaccccgc | ccgcgtctct | accttgaagc | accccgggcg | ctggcactgg | 18540 |
| ccagagcagc | agcagtagta | gcagcagcag | caacgggggtc | ccccgagctc | ctcggggcct | 18600 |
| ccagcccata | gctgtgagct | cctcggcctc | taggcagcgg | ctcgcaactc | cggtcccgcc | 18660 |
| cagggtggat | tgcggccgac | cogtgcocgg | tgcagcctca | ggcgcgcgcc | ttcggaacct | 18720 |
| cccgccccca | cctcccaccg | cccgcctctg | ctcccgctct | cctccccccg | caaccccgct | 18780 |
| cggaagcctgg | ccagggggccc | cgacggcgcg | cgccatgggg | gcgcggggtc | gcacctcccg | 18840 |
| gaccgcgcgc | cctcgagggg | gtggagctgg | gcggaggagg | gaatccgtgc | ggccccctcg | 18900 |
| atgacccggc | cgagccctcc | ctccccgtcg | gtctcagagg | gcctctactc | ctgagaggag | 18960 |
| gagagaaccg | ctgggaaggt | tctttggagg | cgcggcgctg | gtgggatgag | gcgggtgggca | 19020 |
| aaggccgcct | ctcgctcgtg | aagttagggc | caggagcgcg | atcttccgtg | gtctcctggg | 19080 |
| gcgcatctct | tgccccctct | tgctaccogt | cctgccccga | gggtgcctcg | gcggaggttg | 19140 |
| agtccgggtca | tccacctgca | ctgggtgccc | ccaaggatag | gaaggttcag | gcaaccggct | 19200 |
| gcgcctgtct | tgggggcttc | attgctgggg | aaaggcgatg | cagcagacgg | agacaaacct | 19260 |
| tcttccccgg | cgtggggcag | agggcagaat | tgcataaaag | ctgcagactc | ccaggcctgg | 19320 |
| gagacccttt | cggcctcagt | aacatctgtt | tcatgtttta | aaottttgtt | ttcctactcg | 19380 |
| gtgcaaaatt | ggatgagatg | ttaacctttt | tttttttttt | tttttttgaga | tggagtctcc | 19440 |
| ctctgtcgcc | aggtctggag | gcagcgcgcg | gatcttgggt | caactgcaacc | tcgcactccc | 19500 |
| tggttcaagc | gattctcctg | cctcagcctc | ccgagtagct | gggaactacag | gcgcgcgcta | 19560 |
| ccacccccag | ctaacttttg | tatttttagc | agagacgagg | tttcaccatt | tggccaggga | 19620 |
| tgggtctcaat | ctcctgatct | cgtgatccac | cgcctccggc | ctctcaaaagc | gttgggatta | 19680 |
| caggcatgag | ccacgcgcgc | cggcgggaga | tgttaacttt | taaggcaaat | tttttttttt | 19740 |
| tttttttttt | tgagacagag | tttctctctt | gttaccocaga | ctggagtgca | atggcatgat | 19800 |
| ctgggctcac | tgcaacctct | gcctcccaga | ttcaagtgat | tcttctgcct | cagcctcccg | 19860 |
| agtagctggc | attacaggca | ttcgccacca | cgcctgggta | attttgtatt | tttagtagag | 19920 |
| atgggggttc | tccatgttgg | tcgagctggg | ctcgaaactc | cgacctcagg | tgatctgccc | 19980 |
| gcctcggcct | cccaaaagcg | tgggaattaca | ggcgtgagac | accgcaccca | gcctactttt | 20040 |
| aagttaaatct | atttgttttt | gagaattttg | aatgtagtaa | ttttggttagt | gaaagttcga | 20100 |
| gcagtgagag | aaaactacat | tccacatatc | caaaaacaaa | aagtacagaa | agcataggga | 20160 |
| aaagtctccg | tgctctttagc | cctcctcacc | aacaggaaac | caatatgatt | agtttctctt | 20220 |
| ataggctttt | agattatttt | ttcacactca | agacaataca | gacatatatt | tttctcttat | 20280 |
| tacagttttt | ctgcactttg | attttctttt | tttttttggt | cgtttaatac | accttagata | 20340 |
| tcagttcggt | tagagggtcc | tttgtgttct | tatgattatt | atttagagac | aggggtctcc | 20400 |
| tctgtcaccc | acgttagagg | acagttggct | gatcatgcct | cattgcagcc | ttgaaatcct | 20460 |
| gggtctcaag | tatcctccca | cctcagcctc | ctgagtagct | ggaactacag | gcacacggca | 20520 |
| ccagggcccg | ctaaaaattt | taatttttct | gtagacaggg | ggtctcactt | tgtttcccg | 20580 |
| gctgggtcca | aaotctcgtt | cttggccagg | cgcagtgtct | catgcctgta | atccccgac | 20640 |
| tttggggagg | cgaggcgggc | agatcactgg | aggtcaggag | ttcaagacca | gtctggccaa | 20700 |
| catgggtgaa | ccccatctct | actaaaaata | caaaaaattag | cgggcatggg | tsgtgagcgc | 20760 |
| ctgtagttcc | agctactttg | gaggctgagg | caggaaaaac | gcttgaactc | agaaggtgga | 20820 |
| gggtgcagcg | agccgagatc | atgccattgc | actccagcct | gggcaacaag | agcgaaactc | 20880 |
| cgtctcaaaa | aataaaaaata | aaaaataaaa | gaactcctga | ctttaagtga | tctctctgcc | 20940 |
| tcagcttctc | aaatcgctgg | aattacagga | gtgagtcacc | acagctgtcc | agctacgaga | 21000 |
| ttattactta | ttattactac | tttggatttt | caaatcaact | tcattaaagt | ataatttaca | 21060 |
| cacaataaaa | tgcacttatt | tttaagtggc | agtaagatga | gtttcgataa | gtgtatataa | 21120 |
| ctacataaag | atcactataa | tgcagacaca | ttccctcact | cacagaagaa | gcctctgtgc | 21180 |
| cttccagcca | aactttccca | ctccaacccc | cagacagcca | ctgatctgtt | gttctctgtc | 21240 |
| tatagataag | tttgcctgtg | tgtagaattt | catataaaatg | gaatcatgcy | gcatgcactc | 21300 |
| ttctgtgtct | ggcttccctc | cctctttccg | atgttttttg | gattcattta | cactattttg | 21360 |
| cattatcaata | gtttgttccct | tcgtatttgt | gaatagtgtt | cgggtggtttg | agggaaaccac | 21420 |
| agtttctcta | ctcacagctg | caccataagg | gttaggggct | gttaggggct | cttataattg | 21480 |
| gaactatatt | tgcacagaga | gagagagagg | aagaaagagg | gagagagata | tttatctatg | 21540 |
| caatctgctc | acgtgattat | ggagggccaaa | aagttcccca | atctgcactc | tgcagagctg | 21600 |
| agaacgagga | aagccagtg | tgtgattcag | tttgagttca | aaggcctgag | aaccaggagc | 21660 |
| accagtaggc | aggtggctcg | agctcagaac | aagttgggga | caggaaagca | gagcagcacc | 21720 |
| ccagagcagc | ccctcagcga | cacctcttca | gtaaagcaga | gctggaagca | ggcgggctgg | 21780 |
| cttcagtggt | gatgtcaggt | acagaaggca | gctcagggag | ctactctggc | gttcttgcct | 21840 |
| actgggtattc | ttactctgaa | ctggccaact | cctacttaaa | ctgcaggcca | tggctttaat | 21900 |

73/154

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gtctctgtcat | tcagaggctg | tcccttacc | aaagccaggt | tagcatcccc | tgactgacac | 21960 |
| ttctccctgc | aacacgtttc | agaaggccct | gtagtcgtcc | acttccctgt | ctctctcccc | 22020 |
| aagctcctga | gctccatgtg | gtctgggaat | atgtgtgttg | ctcacttcct | agcacagtca | 22080 |
| gtgctaataa | ctgactgtag | aggggacaca | gtcgaaaagc | cacatgggga | tcagagtcac | 22140 |
| ccttacacag | ttgacacctc | ccaaaccag | atgagctgtg | tccaagtgca | ggtcagagga | 22200 |
| atttttctg | gaagtctctg | agaaagggtt | tatttacatt | ttgagggtgc | aggggaggag | 22260 |
| atgaggccat | caaaccaaag | ctgaggaaga | gggatccctag | gatgcaccga | gcagctccgg | 22320 |
| gggcgcctga | cagcacctgg | gaaagatggc | ttctccactg | gcttgttggt | gtcacccctc | 22380 |
| agaggggcat | caggaaatgt | cctgggaacc | aggcaacca | gtgagcatta | acccttagaa | 22440 |
| gtgcttggca | tgggtgacac | ccaccatctg | taaacacgac | ttctcccaag | gagtgacgca | 22500 |
| gaacaggatg | tctgagggag | gcactccgac | tccagccttc | agagatcgcc | aggggtggc | 22560 |
| ctgggtgacga | caggctgatg | cttgggtgcc | ccagaaaagg | tcactgtgtg | gaatgggggc | 22620 |
| ccaaagcca | acgcttcatc | cctgacagcc | tgggtgcatt | agaggggaac | tttttgtccc | 22680 |
| ttggcaagg | gggtggaatt | tcaggttcat | agggcaagg | tatttttagct | ttaatagata | 22740 |
| ttgtcaaaaca | gttttccaaa | gtcattgtac | acactctgtg | attctactta | tgtaaagtgt | 22800 |
| aaaaaacagg | aaatcaaatc | tatggtgttc | gaagtcaaga | cagtagttac | ccttgtgggg | 22860 |
| gctgcaactg | gtacagagt | taagggggga | ctgtaggatg | gtctatttct | tgactctgggt | 22920 |
| gtgttcgctt | ttggaaaagt | ccttgagttg | catttataat | gtgtgaactt | ttctgtatgt | 22980 |
| tacactttaa | ttgaatgtac | aaaaagtctc | aggaggcctc | agaccactgg | aagcggacac | 23040 |
| aactaacccc | tctgagagcc | tccaatccaa | gatggacata | tgtccctctg | gaagtatgca | 23100 |
| gaagcagggtg | aagactccta | agccggatat | tcccaaatcc | ccccagtagc | cgcagcttca | 23160 |
| gcagctgctt | atggctcctc | ctacacccctc | tcttccccag | acagcccca | aacatctggc | 23220 |
| tgcatttgac | ttgctctctc | cctgtccac | ctctgggatt | agtcctatgt | ctccaccctc | 23280 |
| ccactgtca | gcaatgtaga | caagacaaac | gcttagttca | cgtgccacc | tactgcgtgc | 23340 |
| catgcacggg | gctggctcatt | gtgggtggca | aatgtgagca | acacacgaag | ctccaaggag | 23400 |
| cagaaaggga | cacaaatcac | ttcagcgtaa | ggtaatttgt | gataaatgtc | atgttaactg | 23460 |
| cagcccctgg | ccccctccta | cagatgggtg | ctaagaataa | accccactaa | catgtgactc | 23520 |
| ctctgttcta | gccagctgt | ttgggttgca | agaaagagac | tcactccagt | tgcg | 23574 |

<210> 7

<211> 65

<212> DNA

<213> Homo sapiens

<400> 7

| | | | | | | |
|-----------|------------|------------|------------|------------|------------|----|
| agtctcgtg | ctcttagccc | tcttcaccaa | caggaaacca | atatgattag | tttctttcat | 60 |
| aggct | | | | | | 65 |

<210> 8

<211> 656

<212> DNA

<213> Homo sapiens

<400> 8

| | | | | | | |
|-------------|------------|------------|------------|------------|------------|-----|
| gtgcagcctc | aggccgcgc | cttcggacct | tcccgcctcc | acctccacc | gcccgcctc | 60 |
| gtccccct | ccccctcccg | ccaaacccgc | tcggagcctg | gccagggg | ccgacggcg | 120 |
| gcgcacatggg | ggagccgggt | cgcactccc | ggacgcgc | ccctcgagg | ggtggagctg | 180 |
| ggcggaggag | ggaatccgtg | cggccctc | gatgaccgc | ccgagccgtc | cctccccgc | 240 |
| ggtctcagag | ggcctctact | cctgagagga | ggagagaaac | gctgggaagg | ttcttggagg | 300 |
| accgcggcgt | gggtggatga | ggcggtgggc | aaaggccgc | tctcgtctg | gaagttggcc | 360 |
| ccaggagcgc | gatctccgt | ggtctcctgg | ggccgatctc | tgtccctcc | ttgtaccgc | 420 |
| tcctgcccc | aggggtccct | ggcggaggtt | gagtcgggtc | atccacctgc | actgggtgcc | 480 |
| cccaaggata | ggaaggttca | ggcaaccggc | tgcgcgtgtc | ttgggggctt | catgtctggg | 540 |
| caaaggcgat | gcagcgagac | gagacaaact | ttctccctg | gcggtggcca | gagggcagaa | 600 |
| ttgcataaaa | gctgcagact | cccaggcctg | ggagacctt | tcggcctcag | taacat | 656 |

<210> 9
<211> 177
<212> DNA
<213> Homo sapiens

<400> 9
cgggcacggg tcggccgcaa tccagcctgg gcggagccgg agttgcgagc cgctgcctag 60
aggccgagga gctcacagct atgggctgga ggccccggag agctcggggg accccgttgc 120
tgctgctgct actactgctg ctgctctggc cagtgccagg cgccgggggtg cttcaag 177

<210> 10
<211> 80
<212> DNA
<213> Homo sapiens

<400> 10
gacatatccc tgggcagcca gtcaccccg c actgggtcct ggatggacaa ccctggcgca 60
ccgtcagcct ggaggagccg 80

<210> 11
<211> 77
<212> DNA
<213> Homo sapiens

<400> 11
gtctcgaagc cagacatggg gctgggtggc ctggaggctg aaggccagga gtcctgctt 60
gagctggaga agaacca 77

<210> 12
<211> 79
<212> DNA
<213> Homo sapiens

<400> 12
caggctgctg gccccaggat acatagaaac ccactacggc ccagatgggc agccagtgg 60
gctggccccc aaccacacg 79

<210> 13
<211> 119
<212> DNA
<213> Homo sapiens

<400> 13
caggctgctg gccccaggat acatagaaac ccactacggc ccagatgggc agccagtgg 60
gctggccccc aaccacacg tgagatgctt ccatgggctc tgggatgcac cgccagagg 119

<210> 14
<211> 77
<212> DNA
<213> Homo sapiens

<400> 14

75/154

gatcattgcc actaccaagg gcgagtaagg ggcttccccg actcctgggt agtcctctgc 60
 acctgctctg ggatgag 77

<210> 15
 <211> 190
 <212> DNA
 <213> Homo sapiens

<400> 15
 tggcctgac accctcagca ggaatgccag ctattatctg cgtccctggc cccccgggg 60
 ctccaaggac ttctcaaccc acgagatctt tcggatggag cagctgctca cctggaaagg 120
 aacctgtggc cacagggatc ctgggaacaa agcgggcatg accagccttc ctggtggtcc 180
 ccagagcagg 190

<210> 16
 <211> 66
 <212> DNA
 <213> Homo sapiens

<400> 16
 ggccaggcgag aagcgcgcag gacccggaag tacctggaac tgtacattgt ggagaccac 60
 accctg 66

<210> 17
 <211> 72
 <212> DNA
 <213> Homo sapiens

<400> 17
 ttcttgactc ggcaccgaaa ctgaaccac accaaacagc gtctcctgga agtcgccaac 60
 tacgtggacc ag 72

<210> 18
 <211> 167
 <212> DNA
 <213> Homo sapiens

<400> 18
 cttctcagga ctctggacat tcaggtggcg ctgaccggcc tggagtggtg gaccgagcgg 60
 gaccgcagcc gcgtcacgca ggacgccaac gccacgctct gggccttctc gcagtggcgc 120
 cgggggctgt gggcgcagcg gcccacgac tccgcgcagc tgctcac 167

<210> 19
 <211> 85
 <212> DNA
 <213> Homo sapiens

<400> 19
 gggccgcgcc ttccagggcg ccacagtggg cctggcgccc gtcgagggca tgtgccgcgc 60
 cgagagctcg ggaggcgtga gcacg 85

<210> 20

76/154

<211> 143
<212> DNA
<213> Homo sapiens

<400> 20
gaccactcgg agctcccat cggcgccgca gccaccatgg cccatgagat cggccacagc 60
ctcgggctca gccacgacct cgacggctgc tgcgtggagg ctgcggccga gtccggaggc 120
tgcgatcatgg ctgcggccac cgg 143

<210> 21
<211> 178
<212> DNA
<213> Homo sapiens

<400> 21
gcacccgttt ccgcgctgtg tcagcgctg cagccgccc cagctgcgg ccttcttcg 60
caaggggggg ggcgcttgcc tctccaatgc cccggacct ggactcccgg tgcggccggc 120
gctctgcggg aacggcttcg tggaaagcgg cgaggagtgt gactgcggcc ctggccag 178

<210> 22
<211> 90
<212> DNA
<213> Homo sapiens

<400> 22
gagtgcgcgg acctctgctg ctttgcctac aactgctgc tgcgccggg ggcccagtcg 60
gccacgggg actgctgctg gcgctgcctg 90

<210> 23
<211> 196
<212> DNA
<213> Homo sapiens

<400> 23
ctgaagccgg ctggagcgct gtgccgccag gccatgggtg actgtgacct cctgagttt 60
tgacggggca cctcctccca ctgtcccca gacgtttacc tactggacgg ctcacctgt 120
gccaggggca gtggctactg ctgggatggc gcattgtcca cgctggagca gcagtgcag 180
cagctctggg ggcctg 196

<210> 24
<211> 107
<212> DNA
<213> Homo sapiens

<400> 24
gctccacccc agctcccag gcctgtttcc aggtggtgaa ctctgcggga gatgctcatg 60
gaaactgcgg ccaggacagc gagggccact tcctgccctg tgcaggg 107

<210> 25
<211> 199
<212> DNA
<213> Homo sapiens

77/154

<400> 25
 ggatgccttg tgtgggaagc tgcagtgccca ggggtggaaag cccagcctgc tcgcaccgca 60
 catgggtgccca gtggactcta ccgttcacct agatggccag gaagtgactt gtcggggagc 120
 cttgggactc cccagtgcgc agctggacct gcttggcctg ggcttggtag agccaggcac 180
 ccagtgtgga cctagaatg 199

<210> 26
 <211> 109
 <212> DNA
 <213> Homo sapiens

<400> 26
 gtttgcaata gcaaccataa ctgccactgt gctccaggct gggctccacc cttctgtgac 60
 aagccaggct ttggtggcac catggacagt ggccctgtgc aggtgaaa 109

<210> 27
 <211> 148
 <212> DNA
 <213> Homo sapiens

<400> 27
 accatgacac attcctgctg gccatgctcc tcagcgtcct gctgctctg ctcccagggg 60
 cgggcctggc ctggtgttgc taccgactcc caggagccca tctgcagcga tgcagctggg 120
 gctgcagaag ggaccctgcy tgcagtgg 148

<210> 28
 <211> 92
 <212> DNA
 <213> Homo sapiens

<400> 28
 ccccaaagat ggcccacaca gggaccaccc cctggggggc gttcacccca tggagttggg 60
 cccacagcc actggacagc cctggccctt gg 92

<210> 29
 <211> 72
 <212> DNA
 <213> Homo sapiens

<400> 29
 accctgagaa ctctcatgag cccagcagcc accctgagaa gcctctgccca gcagtctcgc 60
 ctgaccccca ag 72

<210> 30
 <211> 1031
 <212> DNA
 <213> Homo sapiens

<400> 30
 cagatcaagt ccagatgccca agatcctgcc tctggtgaga ggtagctcct aaaatgaaca 60
 gatttaaaga caggtggcca ctgacagcca ctccaggaac ttgaactgca ggggcagagc 120
 cagtgaatca ccggacctcc agcacctgca ggcagcttgg aagtttcttc cccgagtgga 180
 gcttcgaccc acccactcca ggaaccacga gccacattag aagttcctga gggctggaga 240

78/154

```

acactgctgg gcacactctc cagctcaata aaccatcagt ccacagaagca aagggtcacac 300
agccctcgac ctccctcacc agtggaggct gggtagtgct ggccatccca aaagggctct 360
gtcctgggag tctggtgtgt ctccctacatg caatttccac ggaccagct ctgtggaggg 420
catgactgct ggccagaagc tagtggctct ggggccctat ggttcgactg agtccacact 480
ccctgcagc ctggctggcc tctgcaaaaca aacataatct tggggacctt ccttcctgtt 540
tcttcccacc ctgtcttctc ccctagggtgg ttcttgagcc cccaccccca atcccagtcg 600
tacacctgag gttctggagc tcagaatctg acagcctctc cccattctg tgtgtgtcgg 660
ggggacagag ggaaccattt aagaaaagat accaaagtag aagtcaaaag aaagacatgt 720
tggctatagg cgtgggtggct catgcctata atcccagcac tttgggaagc cggggtagga 780
ggatcaccag aggcagcag gtccacacca gcctgggcaa cacagcaaga caccgcatct 840
acagaaaaat tttaaaatta gctgggcgtg gtggtgtgta cctgtaggcc tagctgctca 900
ggaggctgaa gcaggaggat cacttgagcc tgagttcaac actgcagtga gctatgggtgg 960
caccactgca ctccagcctg ggtgacagag caagaccctg tctctaaaat aaatttttaa 1020
aagacataaa a                                     1031

```

```

<210> 31
<211> 78
<212> DNA
<213> Homo sapiens

```

```

<400> 31
gtgtgccaga gcaggcgctg caggaagaat gccttcagg agcttcagcg ctgcctgact 60
gcctgccaca gccacggg                                     78

```

```

<210> 32
<211> 6
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: polyhistidine tag

```

```

<400> 32
His His His His His His
1                               5

```

```

<210> 33
<211> 8
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: FLAG epitope tag

```

```

<400> 33
Asp Tyr Lys Asp Asp Asp Asp Lys
1                               5

```

```

<210> 34
<211> 22
<212> DNA
<213> Artificial Sequence

```

```

<220>

```


<223> Description of Artificial Sequence: Primer

<400> 34
aactcttgaa atgagaagcg tg 22

<210> 35

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 35
aatatcatgc accatgaccc ac 22

<210> 36

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 36
tggagtaagt attgtaaact at 22

<210> 37

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 37
ggagcttatc ctggattatc ta 22

<210> 38

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 38
agagccacac atccatgtcc tg 22

<210> 39

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 39

aagccactct gtgaattgcc at 22

<210> 40

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 40

gagtagtcgt agtaccagat gg 22

<210> 41

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 41

gtctggcaat ggagcatgaa aa 22

<210> 42

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 42

attagagcac atgaaggaaa gg 22

<210> 43

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 43

aactgcttt gggggacagg ct 22

<210> 44

<211> 22

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 44
cacgacgccacagagccagc tc 22

<210> 45
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 45
aaccaccacg gattcacgct tc 22

<210> 46
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 46
ataaccagat ggctgtgggt ca 22

<210> 47
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 47
atccccgcaa tgaaatagtt ta 22

<210> 48
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 48
gttgagagcc cacttagata at 22

<210> 49
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 49
gcattggggg aagccaggac at 22

<210> 50
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 50
gccactagga ggcaatggca at 22

<210> 51
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 51
cgacggcatc acggccatct gg 22

<210> 52
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 52
tccaggctca ttcattttca tg 22

<210> 53
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 53
tgacatcaac ttctcctttc ct 22

<210> 54
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 54
agttgcagag acctagcctg tc 22

<210> 55
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 55
tctgggagag gacggagctg gc 22

<210> 56
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 56
tgtaggacta tattgctc 18

<210> 57
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 57
cgacatttag gtgacact 18

<210> 58
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: BstXI-linker
adapter

<400> 58
gtcttcacca cgggg

15

<210> 59
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: BstXI-linker
adapter

<400> 59
gtggtgaaga c

11

<210> 60
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 60
Asp Pro Gln Ala Asp Gln Val Gln Met
1 5

<210> 61
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 61
Asp Pro Gln Asp Gln Val Gln Met
1 5

<210> 62
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<221> MOD_RES
<222> (1)..(11)
<223> "Xaa" represents a variable amino acid

<220>
<223> Description of Artificial Sequence: Zn-binding

85/154

consensus sequence

<400> 62

His Glu Xaa Xaa His Xaa Xaa Gly Xaa Xaa His
1 5 10

<210> 63

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 63

ctgcctagag gccgagga

18

<210> 64

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 64

caggagacca cggaagatcg

20

<210> 65

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 65

ttgcctgaac cttcctatcc

20

<210> 66

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 66

cccctgtggt cctcaggtc

19

<210> 67

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 67

gctccacact ctttcttgcc

20

<210> 68

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 68

aggcaggagg aagctgaat

19

<210> 69

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 69

cctaccacac cctccctctt

20

<210> 70

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 70

cctacccttc tgcacctta

19

<210> 71

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 71

aacttccttc tgggagctgg

20

<210> 72

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 72
cacaccctgg tgaggagaga 20

<210> 73
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 73
ccacgaagga ccaccg 16

<210> 74
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 74
ctcacgtggg tgcctctg 18

<210> 75
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 75
ctctacggcc gcagtgac 18

<210> 76
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 76
gtccctccat gcccaatg 18

<210> 77
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 77
cagggttaagt cggtctgc 18

<210> 78
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 78
ctctctctgc cttccccac 19

<210> 79
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 79
tctactgtgg ggaagatggg 20

<210> 80
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 80
cccctctact tcctcccca 19

<210> 81
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 81

gaccttgggg ttcctaattcc

20

<210> 82
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 82
gtgcacctgc tcaggactc

19

<210> 83
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 83
cctggactct tatcacgttg c

21

<210> 84
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 84
ttaccctcca ccattttctcc

20

<210> 85
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 85
gtggagaggg aagggagaag

20

<210> 86
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 86
ccccatgggt tgaatttaca 20

<210> 87
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 87
gcagctaggc ctacaggtac a 21

<210> 88
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 88
accacgccta tagccaacat 20

<210> 89
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 89
aggtgtagca ctgggattgg 20

<210> 90
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 90
ccccaggacc actagcttct 20

<210> 91
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 91

attgagctgg agagtgtgcc

20

<210> 92

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 92

ttcaagttcc tggagtggct

20

<210> 93

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 93

acaaggaccc tctaaacgca

20

<210> 94

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 94

acccttctgt gacaagccag

20

<210> 95

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 95

gtgttgctac cgactccag

20

<210> 96

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 96

cccaggtgca gagagcag

18

<210> 97

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 97

gctctctcttg tccactctcc t

21

<210> 98

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 98

gccacttcct ctgcacaaat

20

<210> 99

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 99

ttctctgtga cctgggtggt

20

<210> 100

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 100

atttgggcca gagatggg

18

<210> 101

<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 101
ggcagaggag caaggtgg 18

<210> 102
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 102
atggcttgga atcatcaagg 20

<210> 103
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 103
tagagagagg aggtgccagc 20

<210> 104
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 104
aaagatggcc cacacagg 18

<210> 105
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 105
agaactctca tgagcccagc 20

<210> 106
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 106

agctctgagc agaaccatc

20

<210> 107
<211> 18
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 107

ctcgagggg tggagctg

18

<210> 108
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 108

gagaggagga gagaaccgct

20

<210> 109
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 109

agtgacttg tggttctggg

20

<210> 110
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 110

tgtcatctgc accctctctg

20

<210> 111

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 111

aagaggagg gtgtggtagg

20

<210> 112

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 112

gtgatcaggc cactagggtg

20

<210> 113

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 113

atacagcatt cccactccca

20

<210> 114

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 114

gaaggcagaa atcccggt

18

<210> 115

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 115
caccagcacc tgcctgtc 18

<210> 116
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 116
gggtcagagg cacccac 17

<210> 117
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 117
gccgtagagc ctcctgtct 19

<210> 118
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 118
gacgaccaaa gaaacgcag 19

<210> 119
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 119
tgagcggaga gggcaagt 18

<210> 120
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 120

aaaccctcac cctgaacctt

20

<210> 121

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 121

aagggtgctc gtgtcctct

19

<210> 122

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 122

ccactcagct ccactcccta

20

<210> 123

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 123

ggattcaaac ggcaaggag

19

<210> 124

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 124

gctgagtcct gagcaggtg

19

<210> 125

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 125

gaaccgcagg agtaggctc

19

<210> 126

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 126

atatggtcag caggagacc

20

<210> 127

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 127

gcatacctggt ctccatgata a

21

<210> 128

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 128

gaggctttga atccaggtcc

20

<210> 129

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 129

cagcaagaca ccgcatctac

20

<210> 130

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 130
gggacagagg gaaccattta 20

<210> 131
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 131
ttccttctctg tttcttccca 20

<210> 132
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 132
gtcctgggag tctggtgtgt 20

<210> 133
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 133
aggaaccag agccacacta 20

<210> 134
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 134
tgcctctggt gagaggtagc 20

<210> 135
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 135
ttcctggatc actggtcctc 20

<210> 136
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 136
ttcgagcagt gagagaaacc t 21

<210> 137
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 137
ctgggagtcg gtagcaaca 19

<210> 138
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 138
aggccactgg aacctcct 18

<210> 139
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 139

gcagcatggt acagggactg

20

<210> 140
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 140
cagctgacca gtggtatgga

20

<210> 141
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 141
tgtcagacat ggccacagag

20

<210> 142
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 142
agggtcctct tagctgccac

20

<210> 143
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 143
aggccttgtc atttctgtg

20

<210> 144
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 144
caaagaacct tggatgtccg 20

<210> 145
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 145
ctcagctccc ttctgctc 19

<210> 146
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 146
ctgtgtgggc catctttg 18

<210> 147
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 147
ggagaaatgg tggagggtaa 20

<210> 148
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 148
aaagccacag cttctcct 19

<210> 149
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 149

aggtttctgg gctcaggta 20

<210> 150

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 150

gtagggtgtgc cagagcagg 19

<210> 151

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 151

tgtggaccta gaatggtgag c 21

<210> 152

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 152

caaagtcaca caacaagcgg 20

<210> 153

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 153

caggatcttg gcatctggac 20

<210> 154

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 154

ctggcttgtc acagaagggt

20

<210> 155

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 155

ctggagcaca gtggcagtta

20

<210> 156

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 156

tttggtcgtc cctcagtttc

20

<210> 157

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 157

cctctcagga gtagaggccc

20

<210> 158

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 158

agcggttctc tcctectctc

20

<210> 159

105/154

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 159
cctctcagga gtagaggccc 20

<210> 160
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 160
atgttactga ggccgaaagg 20

<210> 161
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 161
ccctttccag ccttctcttt 20

<210> 162
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 162
caggactgca aacatcctga 20

<210> 163
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 163
tccttggtgc ttcccata 18

<210> 164
<211> 19
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 164

aggcaggagg aagctgaat

19

<210> 165
<211> 18
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 165

cctcttgccc ctcttgct

18

<210> 166
<211> 21
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 166

cctgaatgtc cagagtcctg a

21

<210> 167
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 167

ggcctcgagt ccagtatatt

20

<210> 168
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 168

agagcctcct gtctctccct

20

<210> 169

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 169

tcgccctcag cttctcag

18

<210> 170

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 170

tcacgtgggt gcctctga

18

<210> 171

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 171

gggttacttc ccctctctgg

20

<210> 172

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 172

ctgggctttc caccctgg

18

<210> 173

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 173
ctgggctttc caccctgg 18

<210> 174
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 174
tccaggtaggt gaactctgc 19

<210> 175
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 175
tagaatggtg agctctgccc 20

<210> 176
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 176
gaccttgggg ttcctaattcc 20

<210> 177
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 177
ccaagcacac ttgagcgtc 19

<210> 178
<211> 18
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 178

agccatgccc tctgcttt

18

<210> 179

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 179

cagcccaagc acacttga

18

<210> 180

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 180

cccatagctg tgagctcctc

20

<210> 181

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 181

aaagcttcag gaccacaaaa

20

<210> 182

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 182

atcttggtcc ctgccattc

19

<210> 183

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 183

gaggagctc tttcccca

18

<210> 184

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 184

ggaccaccag gaaggctg

18

<210> 185

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 185

aaccccagct cccagaag

18

<210> 186

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 186

ctgctcacct ggaaaggaac

20

<210> 187

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 187

actgcaggaa ggcccagag

19

<210> 188

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 188
accgaaactt gaaccacacc 20

<210> 189
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 189
tgagggacga ccaaagaaac 20

<210> 190
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 190
caaagtcaca caacaagcgg 20

<210> 191
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 191
gaacctgagg gcaccaatta 20

<210> 192
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 192
ttggccttag ttaattggtg c 21

<210> 193
<211> 21
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 193
ttggccttag ttaattgggtg c

21

<210> 194
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 194
ctggagcaca gtggcagtta

20

<210> 195
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 195
aggagtaggc tcaggaagca

20

<210> 196
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 196
tgtactggga ggtagagggc

20

<210> 197
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 197

agaggggtgac ttggagcaga

20

<210> 198

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 198

aggcaataac ccactcagga

20

<210> 199

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 199

cccatggggtt gaatttacat a

21

<210> 200

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 200

gcctctgggtg atcctcctac

20

<210> 201

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 201

actcagtcga accatagggc

20

<210> 202

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 202
tgtgtgacct ttgcttctgg 20

<210> 203
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 203
gcatgaagca atgggagaat 20

<210> 204
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 204
actcagtcga accatagggc 20

<210> 205
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 205
gcaggaaggt gtcatggtct 20

<210> 206
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 206
gcaggaaggt gtcatggtct 20

<210> 207
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 207
gggcattgga gaggcaag 18

<210> 208
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 208
tctgcctccc agattcaagt 20

<210> 209
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 209
agaatgcctt ccaggagctt 20

<210> 210
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 210
gtgttgctac cgactcccag 20

<210> 211
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 211
ctgcttcctg agcctactcc 20

<210> 212
<211> 19
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 212

aacaggagggt tccagtggc

19

<210> 213

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 213

agcgagttgt gattgagggt

20

<210> 214

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 214

tgtgcagggt gaaagtatgc

20

<210> 215

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 215

gccacttcct ctgcacaaat

20

<210> 216

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 216

ctgagcccag aaacctgatt

20

<210> 217

<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 217
gtgagtgagg caccaggg 18

<210> 218
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 218
cctagatggc caggaagtga 20

<210> 219
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 219
ccagaaacct gattaggggg 20

<210> 220
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 220
tacctctcac cagaggcagg 20

<210> 221
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 221
gccagaagct agtggtcctg 20

<210> 222
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 222
gcaggcagct tggaagttt 19

<210> 223
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 223
ttatcatgga gaccaggatg c 21

<210> 224
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 224
gacctggatt caaagcctcc 20

<210> 225
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 225
atgttgcta tagcggtgt 20

<210> 226
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 226

ttatcatgga gaccaggatg c

21

<210> 227

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 227

ctgagtggag ggagcagaag

20

<210> 228

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 228

ctgagtggag ggagcagaag

20

<210> 229

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 229

ccatgagatc ggccacag

18

<210> 230

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 230

atttcaaggc tgcaatgagg

20

<210> 231

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 231
actttctttcc atggcctctg 20

<210> 232
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 232
accacccagg tcacagagaa 20

<210> 233
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 233
tccaagacc aggctatgtc 20

<210> 234
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 234
ctggggatga gaagcagc 18

<210> 235
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 235
cttctccctt ccctctccac 20

<210> 236
<211> 20
<212> DNA
<213> Artificial Sequence

121/154

<220>

<223> Description of Artificial Sequence: Primer

<400> 236

atttgtgcag aggaagtggc

20

<210> 237

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 237

catttcctcc aggctctgac

20

<210> 238

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 238

tcagagcctg gaggaaatgt

20

<210> 239

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 239

gttcctggag tgggtgggt

19

<210> 240

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 240

ctgggagtcg gtagcaaca

19

<210> 241

<211> 41

<212> DNA

122/154

<213> Homo sapiens

<400> 241

gccctctgag accgacgggg agggacggct cgggccggtc a 41

<210> 242

<211> 41

<212> DNA

<213> Homo sapiens

<400> 242

caagaacott ccagcgggtt ctctctctct ctcaggagta g 41

<210> 243

<211> 41

<212> DNA

<213> Homo sapiens

<400> 243

caccatctca gctccacact ctttcttgcc caggtctcga a 41

<210> 244

<211> 41

<212> DNA

<213> Homo sapiens

<400> 244

ccaccatctc agctccacac tctttcttgc ccaggtctcg a 41

<210> 245

<211> 41

<212> DNA

<213> Homo sapiens

<400> 245

acaactaagc catcaccaag gctccttct ctagcccaa g 41

<210> 246

<211> 41

<212> DNA

<213> Homo sapiens

<400> 246

tggtgcttcc catattcaca tctcccacaa ctaagccatc a 41

<210> 247

<211> 41

<212> DNA

<213> Homo sapiens

<400> 247

caggatacat agaaaccac tacggcccag atgggcagcc a 41

<210> 248
<211> 41
<212> DNA
<213> Homo sapiens

<400> 248
ccctccaaat cagaagagac aggaattcac aggcctcgag t 41

<210> 249
<211> 41
<212> DNA
<213> Homo sapiens

<400> 249
agctgctcac ctggaaagga acctgtggcc acagggatcc t 41

<210> 250
<211> 41
<212> DNA
<213> Homo sapiens

<400> 250
acttccttct gggagctggg gttgggggtc agggctcaag c 41

<210> 251
<211> 41
<212> DNA
<213> Homo sapiens

<400> 251
ttcctgcagt ggcgccgggg gctgtgggag cagcggcccc a 41

<210> 252
<211> 41
<212> DNA
<213> Homo sapiens

<400> 252
ggttcagggt gaggggtttcg gggagcttgg gagccggcct g 41

<210> 253
<211> 41
<212> DNA
<213> Homo sapiens

<400> 253
cagagaagcg cggggggttg gggactgtcc ctccatgccc a 41

<210> 254
<211> 41

124/154

<212> DNA
<213> Homo sapiens

<400> 254
ccccctctctg ggctctgcgc gtctggcggc tgtagccaag c 41

<210> 255
<211> 41
<212> DNA
<213> Homo sapiens

<400> 255
cagcgcgcgc cagctgcgcg ccttcttccg caaggggggc g 41

<210> 256
<211> 41
<212> DNA
<213> Homo sapiens

<400> 256
agtggcctcc cagtcaagcg aggggggtgga tccctgcccc a 41

<210> 257
<211> 41
<212> DNA
<213> Homo sapiens

<400> 257
tgctggccat gctcctcagc gtctgtctgc ctctgetccc a 41

<210> 258
<211> 41
<212> DNA
<213> Homo sapiens

<400> 258
ctgctgcctc tgctcccagg ggccggcctg gcttggtgtt g 41

<210> 259
<211> 41
<212> DNA
<213> Homo sapiens

<400> 259
gaagtagctt tgaacaggag gttccagtgg cctcccagtc a 41

<210> 260
<211> 41
<212> DNA
<213> Homo sapiens

<400> 260

125/154

gcctctgtct caccagtttt cggcccttg ccaattctctc t 41

<210> 261
<211> 41
<212> DNA
<213> Homo sapiens

<400> 261
acaaatcacc tctgtcaccc ccttgaagtt cccaaatgct g 41

<210> 262
<211> 41
<212> DNA
<213> Homo sapiens

<400> 262
tccataccac tggtcagctg cgggtctggc tgccccctgtg c 41

<210> 263
<211> 41
<212> DNA
<213> Homo sapiens

<400> 263
ggtgctggct gccctgtgc cagggccctg ccttaacca g 41

<210> 264
<211> 41
<212> DNA
<213> Homo sapiens

<400> 264
ggaaatgaca aggccttggg ggatgggatg gggacagtca a 41

<210> 265
<211> 41
<212> DNA
<213> Homo sapiens

<400> 265
agggtcatg cctcctgcct ccttcagat gggcagcacc c 41

<210> 266
<211> 41
<212> DNA
<213> Homo sapiens

<400> 266
gcccccccc agccccaggg tctcctgtg accatattca c 41

<210> 267

126/154

<211> 41
<212> DNA
<213> Homo sapiens

<400> 267
cctgggcgcc gttcacccca tggagttggg cccacagcc a 41

<210> 268
<211> 41
<212> DNA
<213> Homo sapiens

<400> 268
gccccacagc cactggacag ccctggcccc tgggtgagtg a 41

<210> 269
<211> 41
<212> DNA
<213> Homo sapiens

<400> 269
gccctggccc ctgggtgagt gaggcaccag ggggaggtgg a 41

<210> 270
<211> 41
<212> DNA
<213> Homo sapiens

<400> 270
tgcagcctgg ggccccagtc cttaggggac aacatatcct c 41

<210> 271
<211> 41
<212> DNA
<213> Homo sapiens

<400> 271
cactgagtga ggatgggctc tctgccacac agcttgacg c 41

<210> 272
<211> 41
<212> DNA
<213> Homo sapiens

<400> 272
ctggtcctca ctgagtgagg atgggctctc tgccacacag c 41

<210> 273
<211> 41
<212> DNA
<213> Homo sapiens

127/154

<400> 273
atgacctctt gggtatcatg gagaccagga tgctggaagc c 41

<210> 274
<211> 41
<212> DNA
<213> Homo sapiens

<400> 274
agcaagacac cgcattctaca gaaaaatttt aaaattagct g 41

<210> 275
<211> 41
<212> DNA
<213> Homo sapiens

<400> 275
ggaggatcac cagaggccag cagggtccaca ccagcctggg c 41

<210> 276
<211> 41
<212> DNA
<213> Homo sapiens

<400> 276
atcccagcac ttgggaagc cggggttagga ggatcaccag a 41

<210> 277
<211> 41
<212> DNA
<213> Homo sapiens

<400> 277
agcctggctg gcctctgcaa acaaacataa ttttggggac c 41

<210> 278
<211> 41
<212> DNA
<213> Homo sapiens

<400> 278
actgagtcca cactcccctg cagcctggct ggctctgca a 41

<210> 279
<211> 41
<212> DNA
<213> Homo sapiens

<400> 279
tccaggaacc cagagccaca ttagaagttc ctgagggctg g 41

128/154

<210> 280
<211> 41
<212> DNA
<213> Homo sapiens

<400> 280
ttcttccccg agtggagctt cgacccaccc actccaggaa c 41

<210> 281
<211> 41
<212> DNA
<213> Homo sapiens

<400> 281
tcctcattct cagcagatca agtccagatg ccaagatcct g 41

<210> 282
<211> 41
<212> DNA
<213> Homo sapiens

<400> 282
ctgaggacca cacggggtgg tggttggcgg ggtggtggtt g 41

<210> 283
<211> 41
<212> DNA
<213> Homo sapiens

<400> 283
ggctggcagg ccgagcctag atggcagcca gagccccagg c 41

<210> 284
<211> 41
<212> DNA
<213> Homo sapiens

<400> 284
ctttgctctg tcaactcctgc ctcccttggg cgttcacatt c 41

<210> 285
<211> 41
<212> DNA
<213> Homo sapiens

<400> 285
gtgagctctg cccacccgac ccctccttgc cgtttgaatc c 41

<210> 286
<211> 41
<212> DNA
<213> Homo sapiens

<400> 286
tggcgaggtt actcctacac cgggaggagc accgtcgggt c 41

<210> 287
<211> 41
<212> DNA
<213> Homo sapiens

<400> 287
ggctgctcac tattggggcc gcatcgctcc ctgtcccgct t 41

<210> 288
<211> 41
<212> DNA
<213> Homo sapiens

<400> 288
gcgcgcatcgt ccctgtccc gcttggtgtg tgactttgcg c 41

<210> 289
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 289
gcggtcccac ccggtcg 17

<210> 290
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 290
cctcctctct tggcgac 17

<210> 291
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 291
tccacactct ttcttgcc 18

<210> 292
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 292
gctccacact ctttcttgcc 20

<210> 293
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 293
tcaccaaggc tccttcct 18

<210> 294
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 294
cagaagagac aggaattcac a 21

<210> 295
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 295
tggaaaggaa cctgtggcc 19

<210> 296
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 296

gggtttcggg gagcttg

17

<210> 297

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 297

gggttgggg actgtc

16

<210> 298

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 298

ctctgcgct ctggcg

16

<210> 299

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 299

gccgtccctc cccgtcg

17

<210> 300

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 300

tcctcctcta ttggcgaccc

20

<210> 301

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 301
ctccacactt tttcttgccc a 21

<210> 302
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 302
gctccacact ctttcttgc 19

<210> 303
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 303
tcaccaagcc tccttcct 18

<210> 304
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 304
agaagagacg ggaattcac 19

<210> 305
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 305
tggaaggag cctgtgg 17

<210> 306
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 306
aggggtttcgt ggagcttgg 19

<210> 307
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 307
ggggttggag gactgtcc 18

<210> 308
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 308
gctctgcgca tctggcgg 18

<210> 309
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 309
agtcaagcga gggggtgg 18

<210> 310
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 310
cctcagcgtc ctgctg 16

<210> 311
<211> 18
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 311

aacaggaggt tccagtgg

18

<210> 312

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 312

accagttttc ggcccttt

18

<210> 313

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 313

ctgtcacccc cttgaagt

18

<210> 314

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 314

tcagctgcgg tgctgg

16

<210> 315

<211> 15

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 315

gccttggggg atgga

15

<210> 316

<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 316
tcctgcctcc ttccag 16

<210> 317
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 317
actggacagc cctggc 16

<210> 318
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 318
ctgtgtggca gagagccca 19

<210> 319
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 319
aattatgttt gtttcagag gc 22

<210> 320
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 320
gaacttctag tgtggctct 19

<210> 321
<211> 17
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 321

ccaagggagg caggagt

17

<210> 322
<211> 18
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 322

agtcaagcgt gggggtgg

18

<210> 323
<211> 19
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 323

ctcctcagca tcctgctgc

19

<210> 324
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 324

gaacaggagt ttccagtggc

20

<210> 325
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 325

137/154

caccagtttt tggccctttg 20

<210> 326
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 326
ctgtcaccca cttgaagttc 20

<210> 327
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 327
ggtcagctgt ggtgctgg 18

<210> 328
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 328
aggccttggg agatgggat 19

<210> 329
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 329
tcctgccttc ttccag 16

<210> 330
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 330
actggacagt cctggc 16

<210> 331
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 331
tgtggcaggg agccca 16

<210> 332
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 332
attatgtttg cttgcagagg 20

<210> 333
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 333
ggaacttcta atgtggctct g 21

<210> 334
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 334
cccaagggaa gcaggagtga 20

<210> 335
<211> 55
<212> PRT
<213> Homo sapiens

139/154

<400> 335

Cys Cys Phe Ala His Asn Cys Ser Leu Arg Pro Gly Ala Gln Cys Ala
 1 5 10 15

His Gly Asp Cys Cys Val Arg Cys Leu Leu Lys Pro Ala Gly Ala Leu
 20 25 30

Cys Arg Gln Ala Met Gly Asp Cys Asp Leu Pro Glu Phe Cys Thr Gly
 35 40 45

Thr Ser Ser His Cys Pro Pro
 50 55

<210> 336

<211> 11

<212> PRT

<213> Homo sapiens

<400> 336

Thr Met Ala His Glu Ile Gly His Ser Leu Gly
 1 5 10

<210> 337

<211> 86

<212> PRT

<213> Homo sapiens

<400> 337

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu
 1 5 10 15

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln
 20 25 30

Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly
 35 40 45

Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp
 50 55 60

Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu
 65 70 75 80

Leu Glu Lys Asn His Arg
 85

<210> 338

<211> 48

<212> PRT

<213> Homo sapiens

<400> 338

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu
 1 5 10 15

140/154

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln
20 25 30

Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly
35 40 45

<210> 339

<211> 178

<212> PRT

<213> Homo sapiens

<400> 339

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu
1 5 10 15

Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln
20 25 30

Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly
35 40 45

Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp
50 55 60

Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu
65 70 75 80

Leu Glu Lys Asn His Arg Leu Leu Ala Pro Gly Tyr Ile Glu Thr His
85 90 95

Tyr Gly Pro Asp Gly Gln Pro Val Val Leu Ala Pro Asn His Thr Val
100 105 110

Arg Cys Phe His Gly Leu Trp Asp Ala Pro Pro Glu Asp His Cys His
115 120 125

Tyr Gln Gly Arg Val Arg Gly Phe Pro Asp Ser Trp Val Val Leu Cys
130 135 140

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Cys | Ser | Gly | Met | Ser | Gly | Leu | Ile | Thr | Leu | Ser | Arg | Asn | Ala | Ser |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |

Tyr Tyr Leu Arg Pro Trp Pro Pro Arg Gly Ser Lys Asp Phe Ser Thr
165 170 175

His Glu

<210> 340

<211> 113

<212> PRT

<213> Homo sapiens

<400> 340

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu
1 5 10 15

141/154

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln
 20 25 30
 Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly
 35 40 45
 Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp
 50 55 60
 Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu
 65 70 75 80
 Leu Glu Lys Asn His Gly Leu Ile Thr Leu Ser Arg Asn Ala Ser Tyr
 85 90 95
 Tyr Leu Arg Pro Trp Pro Pro Arg Gly Ser Lys Asp Phe Ser Thr His
 100 105 110

Glu

<210> 341
 <211> 165
 <212> PRT
 <213> Homo sapiens

<400> 341
 Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln
 20 25 30
 Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly
 35 40 45
 Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp
 50 55 60
 Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu
 65 70 75 80
 Leu Glu Lys Asn His Arg Leu Leu Ala Pro Gly Tyr Ile Glu Thr His
 85 90 95
 Tyr Gly Pro Asp Gly Gln Pro Val Val Leu Ala Pro Asn His Thr Asp
 100 105 110
 His Cys His Tyr Gln Gly Arg Val Arg Gly Phe Pro Asp Ser Trp Val
 115 120 125
 Val Leu Cys Thr Cys Ser Gly Met Ser Gly Leu Ile Thr Leu Ser Arg
 130 135 140
 Asn Ala Ser Tyr Tyr Leu Arg Pro Trp Pro Pro Arg Gly Ser Lys Asp
 145 150 155 160

142/154

Phe Ser Thr His Glu
165

<210> 342
<211> 168
<212> PRT
<213> Homo sapiens

<400> 342
Leu Ala Pro Gly Tyr Ile Glu Thr His Tyr Gly Pro Asp Gly Gln Pro
1 5 10 15
Val Val Leu Ala Pro Asn His Thr Asp His Cys His Tyr Gln Gly Arg
20 25 30
Val Arg Gly Phe Pro Asp Ser Trp Val Val Leu Cys Thr Cys Ser Gly
35 40 45
Met Ser Gly Leu Ile Thr Leu Ser Arg Asn Ala Ser Tyr Tyr Leu Arg
50 55 60
Pro Trp Pro Pro Arg Gly Ser Lys Asp Phe Ser Thr His Glu Ile Phe
65 70 75 80
Arg Met Glu Gln Leu Leu Thr Trp Lys Gly Thr Cys Gly His Arg Asp
85 90 95
Pro Gly Asn Lys Ala Gly Met Thr Ser Leu Pro Gly Gly Pro Gln Ser
100 105 110
Arg Gly Arg Arg Lys Ala Arg Arg Thr Arg Lys Tyr Leu Glu Leu Tyr
115 120 125
Ile Val Ala Asp His Thr Leu Phe Leu Thr Arg His Arg Asn Leu Asn
130 135 140
His Thr Lys Gln Arg Leu Leu Glu Val Ala Asn Tyr Val Asp Gln Leu
145 150 155 160
Leu Arg Thr Leu Asp Ile Gln Val
165

<210> 343
<211> 167
<212> PRT
<213> Homo sapiens

<400> 343
Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys
1 5 10 15
Gln Gln Leu Trp Gly Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe
20 25 30
Gln Val Val Asn Ser Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp
35 40 45

143/154

Ser Glu Gly His Phe Leu Pro Cys Ala Gly Arg Asp Ala Leu Cys Gly
 50 55 60
 Lys Leu Gln Cys Gln Gly Gly Lys Pro Ser Leu Leu Ala Pro His Met
 65 70 75 80
 Val Pro Val Asp Ser Thr Val His Leu Asp Gly Gln Glu Val Thr Cys
 85 90 95
 Arg Gly Ala Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu
 100 105 110
 Gly Leu Val Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Asn
 115 120 125
 Ser Asn His Asn Cys His Cys Ala Pro Gly Trp Ala Pro Pro Phe Cys
 130 135 140
 Asp Lys Pro Gly Phe Gly Gly Ser Met Asp Ser Gly Pro Val Gln Ala
 145 150 155 160
 Glu Asn His Asp Thr Phe Leu
 165

<210> 344
 <211> 193
 <212> PRT
 <213> Homo sapiens

<400> 344
 Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys
 1 5 10 15
 Gln Gln Leu Trp Gly Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe
 20 25 30
 Gln Val Val Asn Ser Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp
 35 40 45
 Ser Glu Gly His Phe Leu Pro Cys Ala Gly Arg Asp Ala Leu Cys Gly
 50 55 60
 Lys Leu Gln Cys Gln Gly Gly Lys Pro Ser Leu Leu Ala Pro His Met
 65 70 75 80
 Val Pro Val Asp Ser Thr Val His Leu Asp Gly Gln Glu Val Thr Cys
 85 90 95
 Arg Gly Ala Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu
 100 105 110
 Gly Leu Val Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln
 115 120 125
 Ser Arg Arg Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu
 130 135 140

144/154

Thr Ala Cys His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His
 145 150 155 160

Cys Ala Pro Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly
 165 170 175

Gly Ser Met Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe
 180 185 190

Leu

<210> 345

<211> 126

<212> PRT

<213> Homo sapiens

<400> 345

Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys
 1 5 10 15

Gln Gln Leu Trp Gly Pro Asp Gly Gln Glu Val Thr Cys Arg Gly Ala
 20 25 30

Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu Gly Leu Val
 35 40 45

Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln Ser Arg Arg
 50 55 60

Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu Thr Ala Cys
 65 70 75 80

His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His Cys Ala Pro
 85 90 95

Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly Gly Ser Met
 100 105 110

Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe Leu
 115 120 125

<210> 346

<211> 93

<212> PRT

<213> Homo sapiens

<400> 346

Ala Trp Cys Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg Cys Ser
 1 5 10 15

Trp Gly Cys Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp Gly Pro
 20 25 30

His Arg Asp His Pro Leu Gly Gly Val His Pro Met Glu Leu Gly Pro
 35 40 45

145/154

Thr Ala Thr Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser His Glu
50 55 60

Pro Ser Ser His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro Asp Pro
65 70 75 80

Gln Ala Asp Gln Val Gln Met Pro Arg Ser Cys Leu Trp
85 90

<210> 347

<211> 236

<212> PRT

<213> Homo sapiens

<400> 347

Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys
1 5 10 15

Gln Gln Leu Trp Gly Pro Asp Gly Gln Glu Val Thr Cys Arg Gly Ala
20 25 30

Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu Gly Leu Val
35 40 45

Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln Ser Arg Arg
50 55 60

Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu Thr Ala Cys
65 70 75 80

His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His Cys Ala Pro
85 90 95

Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly Gly Ser Met
100 105 110

Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe Leu Leu Ala
115 120 125

Met Leu Leu Ser Val Leu Leu Pro Leu Leu Pro Gly Ala Gly Leu Ala
130 135 140

Trp Cys Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg Cys Ser Trp
145 150 155 160

Gly Cys Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp Gly Pro His
165 170 175

Arg Asp His Pro Leu Gly Gly Val His Pro Met Glu Leu Gly Pro Thr
180 185 190

Ala Thr Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser His Glu Pro
195 200 205

Ser Ser His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro Asp Pro Gln
210 215 220

146/154

Ala Asp Gln Val Gln Met Pro Arg Ser Cys Leu Trp
 225 230 235

<210> 348

<211> 302

<212> PRT

<213> Homo sapiens

<400> 348

Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys
 1 5 10 15

Gln Gln Leu Trp Gly Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe
 20 25 30

Gln Val Val Asn Ser Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp
 35 40 45

Ser Glu Gly His Phe Leu Pro Cys Ala Gly Arg Asp Ala Leu Cys Gly
 50 55 60

Lys Leu Gln Cys Gln Gly Gly Lys Pro Ser Leu Leu Ala Pro His Met
 65 70 75 80

Val Pro Val Asp Ser Thr Val His Leu Asp Gly Gln Glu Val Thr Cys
 85 90 95

Arg Gly Ala Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu
 100 105 110

Gly Leu Val Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln
 115 120 125

Ser Arg Arg Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu
 130 135 140

Thr Ala Cys His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His
 145 150 155 160

Cys Ala Pro Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly
 165 170 175

Gly Ser Met Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe
 180 185 190

Leu Leu Ala Met Leu Leu Ser Val Leu Leu Pro Leu Leu Pro Gly Ala
 195 200 205

Gly Leu Ala Trp Cys Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg
 210 215 220

Cys Ser Trp Gly Cys Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp
 225 230 235 240

Gly Pro His Arg Asp His Pro Leu Gly Gly Val His Pro Met Glu Leu
 245 250 255

147/154

Gly Pro Thr Ala Thr Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser
260 265 270

His Glu Pro Ser Ser His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro
275 280 285

Asp Pro Gln Asp Gln Val Gln Met Pro Arg Ser Cys Leu Trp
290 295 300

<210> 349

<211> 235

<212> PRT

<213> Homo sapiens

<400> 349

Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys
1 5 10 15

Gln Gln Leu Trp Gly Pro Asp Gly Gln Glu Val Thr Cys Arg Gly Ala
20 25 30

Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu Gly Leu Val
35 40 45

Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln Ser Arg Arg
50 55 60

Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu Thr Ala Cys
65 70 75 80

His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His Cys Ala Pro
85 90 95

Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly Gly Ser Met
100 105 110

Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe Leu Leu Ala
115 120 125

Met Leu Leu Ser Val Leu Leu Pro Leu Leu Pro Gly Ala Gly Leu Ala
130 135 140

Trp Cys Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg Cys Ser Trp
145 150 155 160

Gly Cys Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp Gly Pro His
165 170 175

Arg Asp His Pro Leu Gly Gly Val His Pro Met Glu Leu Gly Pro Thr
180 185 190

Ala Thr Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser His Glu Pro
195 200 205

Ser Ser His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro Asp Pro Gln
210 215 220

148/154

Asp Gln Val Gln Met Pro Arg Ser Cys Leu Trp
 225 230 235

<210> 350
 <211> 339
 <212> DNA
 <213> Homo sapiens

<400> 350
 cgggcacggg tcggccgcaa tccagcctgg gcggagccgg agttgcgagc cgctgcctag 60
 agggccgagga gctcacagct atgggctgga ggccccggag agctcggggg accccgttgc 120
 tgctgctgct actactgctg ctgctctggc cagtgccagg cgccgggggtg cttcaaggac 180
 atatccctgg gcagccagtc accccgcact gggtcctgga tggacaaccc tggcgccaccg 240
 tcagcctgga ggagccggtc tcgaagccag acatggggct ggtggccctg gaggtgaag 300
 gccaggagct cctgcttag ctggagaaga accacaggc 339

<210> 351
 <211> 225
 <212> DNA
 <213> Homo sapiens

<400> 351
 cgggcacggg tcggccgcaa tccagcctgg gcggagccgg agttgcgagc cgctgcctag 60
 agggccgagga gctcacagct atgggctgga ggccccggag agctcggggg accccgttgc 120
 tgctgctgct actactgctg ctgctctggc cagtgccagg cgccgggggtg cttcaaggac 180
 atatccctgg gcagccagtc accccgcact gggtcctgga tggac 225

<210> 352
 <211> 562
 <212> DNA
 <213> Homo sapiens

<400> 352
 gcctagaggc cgaggagctc acagctatgg gctggaggcc ccggagagct cgggggaccc 60
 cgttgctgct gctgctacta ctgctgctgc tctggccagt gccaggcgcc ggggtgcttc 120
 aaggacatat ccctggggcag ccagtcaccc cgcactgggt cctggatgga caacctggc 180
 gcaccgtcag cctggaggag ccggctctga agccagacat ggggctggtg gccctggagg 240
 ctgaaggcca ggagctcctg cttgagctgg agaagaacca caggctgctg gccccaggat 300
 acatagaaac ccactacggc ccagatgggc agccagtggg gctggccccc aaccacacgg 360
 tgagatgctt ccatgggctc tgggatgcac cgccagagga tcattgccac taccaagggc 420
 gagtaagggg cttccccgac tcctgggtag tcctctgcac ctgctctggg atgagtggcc 480
 tgatcacctt cagcaggaat gccagctatt atctgcgtcc ctggccaccc cggggctcca 540
 aggacttctc aaccacagag at 562

<210> 353
 <211> 362
 <212> DNA
 <213> Homo sapiens

<400> 353
 gaggccgagg agctcacagc tatgggctgg agggcccgga gagctcgggg gaccccggtt 60
 ctgctgctgc tactactgct gctgctctgg ccagtgccag gcgccggggg gcttcaagga 120
 catatccctg ggcagccagt caccgccgac tgggtcctgg atggacaacc ctggcgccacc 180

149/154

| | | | | | | |
|------------|-------------|------------|------------|-------------|------------|-----|
| gtcagcctgg | aggagccgggt | ctcgaagcca | gacatggggc | tgggtggccct | ggaggctgaa | 240 |
| ggccaggagc | tccctgcttga | gctggagaag | aacctgggcc | tgatcacccct | cagcaggaat | 300 |
| gccagctatt | atctgcgtcc | ctggccaccc | cggggctcca | aggacttctc | aaccacagag | 360 |
| at | | | | | | 362 |

<210> 354

<211> 518

<212> DNA

<213> Homo sapiens

<400> 354

| | | | | | | |
|------------|-------------|------------|------------|-------------|------------|-----|
| gaggccgagg | agctcacagc | tatgggctgg | aggccccgga | gagctcgggg | gaccccgctg | 60 |
| ctgctgctgc | tactactgct | gctgctctgg | ccagtgccag | gcgccgggggt | gcttcaagga | 120 |
| catatccctg | ggcagccagt | caccccgcac | tgggtcctgg | atggacaacc | ctggcgccac | 180 |
| gtcagcctgg | aggagccgggt | ctcgaagcca | gacatggggc | tgggtggccct | ggaggctgaa | 240 |
| ggccaggagc | tccctgcttga | gctggagaag | aaccacaggc | tgctggcccc | aggatacata | 300 |
| gaaacccact | acggcccaga | tgggcagcca | gtggtgctgg | cccccaacca | cacggatcat | 360 |
| tgccactacc | aaggggcagt | aaggggcttc | cccgaactct | gggtagtcct | ctgcaactgc | 420 |
| tctgggatga | gtggcctgat | caccctcagc | aggaatgcc | gctattatct | gcgtccctgg | 480 |
| ccaccccggg | gctccaagga | cttctcaacc | cacgagat | | | 518 |

<210> 355

<211> 506

<212> DNA

<213> Homo sapiens

<400> 355

| | | | | | | |
|------------|------------|-------------|------------|-------------|------------|-----|
| ctggccocag | gatacataga | aaccactac | ggcccagatg | ggcagccagt | ggtgctggcc | 60 |
| cccaaccaca | cggatcattg | ccactaccaa | gggcgagtaa | ggggcttccc | cgactcctgg | 120 |
| gtagtcctct | gcacctgctc | tgggatgagt | ggcctgatca | ccctcagcag | gaatgccagc | 180 |
| tattatctgc | gtccctggcc | accccggggc | tccaaggact | tctcaaccaca | cgagatcttt | 240 |
| cggatggagc | agctgctcac | ctggaaagga | acctgtggcc | acagggatcc | tgggaacaaa | 300 |
| gcgggcatga | ccagccttcc | tgggtggctcc | cagagcaggg | gcaggcgaaa | agcgcgagg | 360 |
| accgggaagt | acctggaact | gtacattgtg | gcagaccaca | ccctgttctt | gactcggcac | 420 |
| cgaacttga | accacaccaa | acagcgtctc | ctggaagtcc | ccaactacgt | ggaccagctt | 480 |
| ctcaggactc | tggacattca | ggtggc | | | | 506 |

<210> 356

<211> 503

<212> DNA

<213> Homo sapiens

<400> 356

| | | | | | | |
|-------------|-------------|------------|------------|------------|------------|-----|
| cagtggtctac | tgctgggatg | gogcatgtcc | cacgctggag | cagcagtgcc | agcagctctg | 60 |
| ggggcctggc | tcacaccag | ctcccagggc | ctgtttccag | gtggtgaact | ctgcgggaga | 120 |
| tgctcatgga | aactgcggcc | aggacagcga | gggccaactc | ctgcctctgt | cagggaggga | 180 |
| tgcccttgt | gggaagtgc | agtcccagg | tggaaagccc | agcctgctcg | caccgcacat | 240 |
| ggtgccagtg | gactctaccg | ttcacctaga | tggccaggaa | gtgacttgtc | ggggagcctt | 300 |
| ggcactcccc | agtccccagc | tggacctgct | tggcctgggc | ctggtagagc | cagggaccca | 360 |
| gtgtggacct | agaatggttt | gcaatatgaa | ccataactgc | cactgtgctc | caggctgggc | 420 |
| tcacaccttc | tgtagacaagc | cagctttgg | tggcagcatg | gacagtgggc | ctgtgcaggc | 480 |
| tgaaaaccat | gacaccttcc | tgc | | | | 503 |

<210> 357

150/154

<211> 581
 <212> DNA
 <213> Homo sapiens

<400> 357
 cagtggctac tgctgggatg ggcgatgtcc cacgctggag cagcagtgcc agcagctctg 60
 ggggcctggc tcccacccag ctcccagggc ctgtttccag gtgggtgaact ctgcgggaga 120
 tgctcatgga aactgcggcc aggacagcga gggccacttc ctgccctgtg cagggaggga 180
 tgccctgtgt gggaaagctgc agtgccaggg tggaaagccc agcctgtctg caccgcacat 240
 ggtgccagtg gactctaccg ttcacctaga tggccaggaa gtgacttgtc ggggagcctt 300
 ggcactcccc agtgcccagc tggacctgct tggcctgggc ctggtagagc caggcaccca 360
 gtgtggacct agaattggtg gccagagcag gcgctgcagg aagaatgcct tccaggagct 420
 tcacgcctgc ctgactgcct gccacagcca cggggtttgc aatagcaacc ataactgcca 480
 ctgtgtctcca ggctgggctc cacccttctg tgacaagcca ggctttggtg gcagcatgga 540
 cagtggccct gtgcaggctg aaaccatga caccctcctg c 581

<210> 358
 <211> 380
 <212> DNA
 <213> Homo sapiens

<400> 358
 cagtggctac tgctgggatg ggcgatgtcc cacgctggag cagcagtgcc agcagctctg 60
 ggggcctgat ggcaggaag tgacttgtcg gggagccttg gcactcccca gtcccagct 120
 ggacctgctt ggctgggccc tggtagagcc aggcacccag tgtggacctt gaattggtgt 180
 ccagagcagg cgctgcagga agaattgcctt ccaggagctt cagcgtgcc tgaactgcctg 240
 ccacagccac ggggttttga atagcaacca taactgccac tgtgtccag gctgggctcc 300
 acccttctgt gacaagccag gctttggtgg cagcatggac agtggccctg tgcaggctga 360
 aaaccatgac accttctctg 380

<210> 359
 <211> 324
 <212> DNA
 <213> Homo sapiens

<400> 359
 ggcttggtgt tgctaccgac tcccaggagc ccatctgcag cgatgcagct ggggctgcag 60
 aagggaacct gcgtgcagtg gccccaaaaga tggccacac agggaccacc ccctgggcgg 120
 cgttcacccc atggagttag gccccacagc cactggacag ccttgccccc tggacctga 180
 gaactctcat gagcccagca gccaccctga gaagcctctg ccagcagctc cgctgaccc 240
 ccaagcagat caagtcaga tgccaagatc ctgectctgg tgagaggtag ctctaaaaat 300
 gaacagattt aaagacaggt ggcc 324

<210> 360
 <211> 753
 <212> DNA
 <213> Homo sapiens

<400> 360
 cagtggctac tgctgggatg ggcgatgtcc cacgctggag cagcagtgcc agcagctctg 60
 ggggcctgat ggcaggaag tgacttgtcg gggagccttg gcactcccca gtcccagct 120
 ggacctgctt ggctgggccc tggtagagcc aggcacccag tgtggacctt gaattggtgt 180
 ccagagcagg cgctgcagga agaattgcctt ccaggagctt cagcgtgcc tgaactgcctg 240
 ccacagccac ggggttttga atagcaacca taactgccac tgtgtccag gctgggctcc 300
 acccttctgt gacaagccag gctttggtgg cagcatggac agtggccctg tgcaggctga 360

151/154

| | | | | | | |
|-------------|-------------|------------|-------------|-------------|-------------|-----|
| aaaccatgac | accttctctgc | tggccatgct | cctcagcgtc | ctgctgcctc | tgctcccagg | 420 |
| ggcggcctg | gcctgggtgt | gctacogact | cccaggagcc | catctgcagc | gatgcagctg | 480 |
| gggtctgcaga | agggaccctg | cgtgcagtg | ccccaaagat | ggccccacaca | gggaccaccc | 540 |
| cctggggcggc | gttcacccca | tggagttggg | ccccacagcc | actggacagc | cctggccctc | 600 |
| ggacccctgag | aactctcatg | agccacagc | ccaccctgag | aagcctctgc | cagcagctctc | 660 |
| gcctgacccc | caagcagatc | aagtcagat | gcccaagatcc | tgcctctggg | gagaggtagc | 720 |
| tcctaaaatg | aacagattta | aagacaggtg | gcc | | | 753 |

<210> 361

<211> 1154

<212> DNA

<213> Homo sapiens

<400> 361

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| cagtggtctac | tgctgggatg | gcgcatgtcc | cacgctggag | cagcagtgcc | agcagctctg | 60 |
| ggggcctggc | tcccaccagc | ctcccaggag | ctgtttccag | gtggtgaact | ctgcggggaga | 120 |
| tgctcatgga | aactgcggcc | aggacagcga | gggccacttc | ctgccctgtg | caggggaggga | 180 |
| tgcctgtgt | gggaagctgc | agtgcacagg | tggaaagccc | agcctgctcg | caccgcacat | 240 |
| ggtgccagtg | gactctaccg | ttcacctaga | tggccaggaa | gtgacttgtc | ggggagcctt | 300 |
| ggcactcccc | agtgccccag | tggacctgct | tggcctgggc | ctggtagagc | caggcaccga | 360 |
| gtgtggacct | agaatgggtg | gccagagcag | gcgctgcagg | aagaatgcct | tccaggagct | 420 |
| tcagcgctgc | ctgactgcct | gccacagcca | cggggtttgc | aatagcaacc | ataactgcca | 480 |
| ctgtgtccca | ggctgggctc | cacccttctg | tgacaagcca | ggctttgggtg | gcagcatgga | 540 |
| cagtggccct | gtgcaggctg | aaaaccatga | caccttccctg | ctggccatgc | tcctcagcgt | 600 |
| cctgtgcct | ctgtccccag | gggccggcct | ggcctgggtg | tgctaccgac | tcccaggagc | 660 |
| ccatctgcag | cgatgcagct | ggggctgcag | aagggaccct | gcgtgcagtg | gccccaaaga | 720 |
| tggcccaac | agggaccacc | cctcgggcgg | cgttcacccc | atggagtggg | gccccacagc | 780 |
| cactggacag | cctcggccccc | tggacccctga | gaactctcat | gagccacagc | gccaccctga | 840 |
| gaagcctctg | ccagcagctc | cgcctgaccc | ccaagatcaa | gtccagatgc | caagatccctg | 900 |
| cctctgggtga | gaggtagctc | ctaaaaatgaa | cagatttaaa | gacagggtggc | cactgcacagc | 960 |
| cactccaggga | acttgaactg | cagggggcaga | gccagtgaa | caccggacct | ccagcacctc | 1020 |
| caggcagctt | gggaagttct | tcccgcagtg | gagcttcgac | ccaccactc | cagggaaccca | 1080 |
| gagccacatt | agaagttcct | gagggtctga | gaacactgct | gggcacactc | tccagctcaa | 1140 |
| taaacatca | gtcc | | | | | 1154 |

<210> 362

<211> 953

<212> DNA

<213> Homo sapiens

<400> 362

| | | | | | | |
|-------------|-------------|------------|------------|-------------|-------------|-----|
| cagtggtctac | tgctgggatg | gcgcatgtcc | cacgctggag | cagcagtgcc | agcagctctg | 60 |
| ggggcctgat | ggccaggaa | tgacttgtcg | gggagccttg | gcactcccca | gtgccagct | 120 |
| ggacctgctt | ggcctgggccc | tggtagagcc | aggcaccac | tgtggacctga | gaatgggtg | 180 |
| ccagagcagg | cgctgcaggga | agaatgcctt | ccaggagctt | cagcgtgcct | tgactgcctg | 240 |
| ccacagccac | ggggttttga | atagcaacca | taactgccac | tgtgtccag | gctgggctcc | 300 |
| acccttctgt | gacaagccag | gctttgggtg | cagcatggac | agtggccctg | tgcaggctga | 360 |
| aaaccatgac | accttccctg | tggccatgct | cctcagcgtc | ctgctgcctc | tgctcccagg | 420 |
| ggccggcctg | gcctgggtgt | gctaccgact | cccaggagcc | catctgcagc | gatgcagctg | 480 |
| gggtctgcaga | agggaccctg | cgtgcagtg | ccccaaagat | ggccccacaca | gggaccaccc | 540 |
| cctggggcggc | gttcacccca | tggagttggg | ccccacagcc | actggacagc | cctggccctc | 600 |
| ggacccctgag | aactctcatg | agcccagcag | ccaccctgag | aagcctctgc | cagcagctctc | 660 |
| gcctgacccc | caagatcaag | tccagatgcc | aagatcctgc | ctctggtgag | aggtagctcc | 720 |
| taaaatgaac | agatttaaa | acaggtggcc | actgacagcc | actccaggaa | cttgaactgc | 780 |
| agggggcagag | ccagtgaa | accggacctc | cagcacctgc | aggcagcttg | gaagtttctt | 840 |
| ccccgagtg | agcttcgacc | caccacactc | aggaacccag | agccacatta | gaagttctctg | 900 |

152/154

agggctggag aacactgctg ggcacactct ccagctcaat aaaccatcag tcc

953

<210> 363

<211> 812

<212> PRT

<213> Homo sapiens

<400> 363

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Trp | Arg | Pro | Arg | Arg | Ala | Arg | Gly | Thr | Pro | Leu | Leu | Leu | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Leu | Leu | Leu | Leu | Leu | Trp | Pro | Val | Pro | Gly | Ala | Gly | Val | Leu | Gln |
| | | | 20 | | | | | 25 | | | | | 30 | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | His | Ile | Pro | Gly | Gln | Pro | Val | Thr | Pro | His | Trp | Val | Leu | Asp | Gly |
| | | 35 | | | | | 40 | | | | | 45 | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gln | Pro | Trp | Arg | Thr | Val | Ser | Leu | Glu | Glu | Pro | Val | Ser | Lys | Pro | Asp |
| | 50 | | | | | 55 | | | | | 60 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Leu | Val | Ala | Leu | Glu | Ala | Glu | Gly | Gln | Glu | Leu | Leu | Leu | Glu |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Glu | Lys | Asn | His | Arg | Leu | Leu | Ala | Pro | Gly | Tyr | Ile | Glu | Thr | His |
| | | | | 85 | | | | | 90 | | | | | 95 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Tyr | Gly | Pro | Asp | Gly | Gln | Pro | Val | Val | Leu | Ala | Pro | Asn | His | Thr | Asp |
| | | | 100 | | | | | 105 | | | | | 110 | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| His | Cys | His | Tyr | Gln | Gly | Arg | Val | Arg | Gly | Phe | Pro | Asp | Ser | Trp | Val |
| | | 115 | | | | | 120 | | | | | 125 | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Leu | Cys | Thr | Cys | Ser | Gly | Met | Ser | Gly | Leu | Ile | Thr | Leu | Ser | Arg |
| | 130 | | | | | 135 | | | | | 140 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Ala | Ser | Tyr | Tyr | Leu | Arg | Pro | Trp | Pro | Pro | Arg | Gly | Ser | Lys | Asp |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Ser | Thr | His | Glu | Ile | Phe | Arg | Met | Glu | Gln | Leu | Leu | Thr | Trp | Lys |
| | | | 165 | | | | | | 170 | | | | | 175 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Thr | Cys | Gly | His | Arg | Asp | Pro | Gly | Asn | Lys | Ala | Gly | Met | Thr | Ser |
| | | | 180 | | | | | 185 | | | | | 190 | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Pro | Gly | Gly | Pro | Gln | Ser | Arg | Gly | Arg | Arg | Glu | Ala | Arg | Arg | Thr |
| | | 195 | | | | | 200 | | | | | 205 | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Lys | Tyr | Leu | Glu | Leu | Tyr | Ile | Val | Ala | Asp | His | Thr | Leu | Phe | Leu |
| | 210 | | | | | 215 | | | | | 220 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Arg | His | Arg | Asn | Leu | Asn | His | Thr | Lys | Gln | Arg | Leu | Leu | Glu | Val |
| 225 | | | | | 230 | | | | | 235 | | | | 240 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Asn | Tyr | Val | Asp | Gln | Leu | Leu | Arg | Thr | Leu | Asp | Ile | Gln | Val | Ala |
| | | | | 245 | | | | | 250 | | | | | 255 | |

153/154

Leu Thr Gly Leu Glu Val Trp Thr Glu Arg Asp Arg Ser Arg Val Thr
 260 265 270

Gln Asp Ala Asn Ala Thr Leu Trp Ala Phe Leu Gln Trp Arg Arg Gly
 275 280 285

Leu Trp Ala Gln Arg Pro His Asp Ser Ala Gln Leu Leu Thr Gly Arg
 290 295 300

Ala Phe Gln Gly Ala Thr Val Gly Leu Ala Pro Val Glu Gly Met Cys
 305 310 315 320

Arg Ala Glu Ser Ser Gly Gly Val Ser Thr Asp His Ser Glu Leu Pro
 325 330 335

Ile Gly Ala Ala Ala Thr Met Ala His Glu Ile Gly His Ser Leu Gly
 340 345 350

Leu Ser His Asp Pro Asp Gly Cys Cys Val Glu Ala Ala Ala Glu Ser
 355 360 365

Gly Gly Cys Val Met Ala Ala Thr Gly His Pro Phe Pro Arg Val
 370 375 380

Phe Ser Ala Cys Ser Arg Arg Gln Leu Arg Ala Phe Phe Arg Lys Gly
 385 390 395 400

Gly Gly Ala Cys Leu Ser Asn Ala Pro Asp Pro Gly Leu Pro Val Pro
 405 410 415

Pro Ala Leu Cys Gly Asn Gly Phe Val Glu Ala Gly Glu Glu Cys Asp
 420 425 430

Cys Gly Pro Gly Gln Glu Cys Arg Asp Leu Cys Cys Phe Ala His Asn
 435 440 445

Cys Ser Leu Arg Pro Gly Ala Gln Cys Ala His Gly Asp Cys Cys Val
 450 455 460

Arg Cys Leu Leu Lys Pro Ala Gly Ala Leu Cys Arg Gln Ala Met Gly
 465 470 475 480

Asp Cys Asp Leu Pro Glu Phe Cys Thr Gly Thr Ser Ser His Cys Pro
 485 490 495

Pro Asp Val Tyr Leu Leu Asp Gly Ser Pro Cys Ala Arg Gly Ser Gly
 500 505 510

Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys Gln Gln
 515 520 525

Leu Trp Gly Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe Gln Val
 530 535 540

Val Asn Ser Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp Ser Glu
 545 550 555 560

154/154

Gly His Phe Leu Pro Cys Ala Gly Arg Asp Ala Leu Cys Gly Lys Leu
 565 570

Gln Cys Gln Gly Gly Lys Pro Ser Leu Leu Ala Pro His Met Val Pro
 580 585 590

Val Asp Ser Thr Val His Leu Asp Gly Gln Glu Val Thr Cys Arg Gly
 595 600 605

Ala Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu Gly Leu
 610 615 620

Val Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln Ser Arg
 625 630 635 640

Arg Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu Thr Ala
 645 650 655

Cys His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His Cys Ala
 660 665 670

Pro Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly Gly Ser
 675 680 685

Met Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe Leu Leu
 690 695 700

Ala Met Leu Leu Ser Val Leu Leu Pro Leu Leu Pro Gly Ala Gly Leu
 705 710 715 720

Ala Trp Cys Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg Cys Ser
 725 730 735

Trp Gly Cys Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp Gly Pro
 740 745 750

His Arg Asp His Pro Leu Gly Gly Val His Pro Met Glu Leu Gly Pro
 755 760 765

Thr Ala Thr Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser His Glu
 770 775 780

Pro Ser Ser His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro Asp Pro
 785 790 795 800

Gln Asp Gln Val Gln Met Pro Arg Ser Cys Leu Trp
 805 810